

گزارش پروژه‌ی کارشناسی

پارسا عالیان - ۹۵۱۰۹۵۲۹

۱ هدف کلی

هدف کلی این پروژه، ایجاد یک سیستم آموزش معاملات بورسی است. در این سیستم از المان‌های مختلفی استفاده می‌شود که در ادامه‌ی گزارش توضیح داده خواهد شد.

۲ مدل‌های یادگیری

به‌طور سنتی، یادگیری تنها در محیط کلاس‌های حضوری امکان‌پذیر بوده است. با گسترش محیط اینترنت، در سال‌های اخیر شاهد رشد سریع محیط‌های یادگیری آنلاین به فرم Massive Open Online Courses (MOOC) بوده‌ایم. در دوران پاندمی Covid-19 نیز علاوه بر این محیط‌ها، محیط‌های دانشگاهی به‌صورت آموزش آنلاین درآمده‌اند. در چنین شرایطی، یادگیری بدون نظارت مستقیم استاد صورت می‌گیرد، یا حتی بعضاً یک شخص حقیقی به‌عنوان استاد در محیط وجود ندارد، مانند سایت‌های آموزش برنامه‌نویسی. حجم افرادی که از این محیط‌ها استفاده می‌کنند نیز از عواملی است که به پیچیدگی این شرایط می‌افزاید. بنابراین ایجاد مدل‌هایی برای یادگیری بدین شکل ضروری است. در سروی Activating learning at scale، روش‌های مختلفی که در یادگیری آنلاین مورد استفاده قرار گرفته‌اند بررسی شده است. این مقاله این روش‌ها را به ده‌دسته تقسیم می‌کند که به شرح زیر هستند. [خلاصه‌ی نکات مقاله]

با توجه به نکات مطرح شده در این مقاله، نیازهای ما از پارامترهای مختلف به شکل زیر است.

۱. محیط: محیط یادگیری بدون ناظر و در شرایط غیرآزمایشگاهی است. همچنین موضوع مورد بررسی محدود است. بنابراین سامانه‌ی مورد استفاده یک Intelligent Tutoring System (ITS) خواهد بود.
۲. مشوق‌ها: از آنجایی که افراد حاضر در سیستم در سطوح دموگرافیک مختلف قرار دارند، مشوق نمره یا حضور در کورس‌ها و کلاس‌ها به احتمال زیاد مشوق‌های مناسبی نیست. برای شرکت در یادگیری، توکنی به محصل داده خواهد شد که می‌تواند دارای ارزش مالی باشد (در ادامه توضیح داده خواهد شد).
۳. خروجی: هدف یادگیری در نهایت این است که افراد حاضر در آن، بتوانند بازدهی خوبی در بازارهای مالی نشان دهند. تعریف بازدهی و معیارهای مورد نظر را در ادامه ارائه خواهیم داد.
۴. نوع یادگیری: با توجه به انواع یادگیری توضیح داده شده، مدل استفاده‌شده در این سیستم به چند بخش مختلف تقسیم می‌شود. استفاده از مدل‌های یادگیری همکارانه، بازی و شبیه‌سازی و در نهایت پرسش و پاسخ در این سامانه بررسی می‌شود.

۳ شرح سناریو

فرمت کلی سیستم یادگیری به این صورت است:

هدف بازی به وجود آوردن توانایی پیش‌بینی و خرید و فروش سهام در اشخاص مختلف است. در ابتدا، هر شخصی درجه‌ای از تخصص دارد که فرض کنید آن را ϵ بنامیم. سنجش میزان ϵ را بر اساس تعدادی سوال از شخص انجام می‌دهیم. هر کدام از سوالات فرمتی به صورت چندگزینه‌ای یا میزان عددی دارند که شخص باید آن را بر اساس داده‌های ورودی مسئله تخمین بزند و انتخاب کند. برای مثال، ورودی نموداری از قیمت یک سهم (به صورت ناشناس و بدون اسم) است که شخص با توجه به حرکت آن، باید تخمین بزند که در یک بازه‌ی زمانی مشخص رشد سهم چه قدر خواهد بود. انواع سوالات قابل تغییر به شکل‌های دیگر از جمله درگیر کردن داده‌های بنیادین یا شاخص‌های تکنیکال است.

پس از انتخاب گزینه یا تخمین مقدار، بر اساس درستی یا میزان صحت جواب شخص، مقداری جایزه به او تعلق می‌گیرد. همچنین با توجه به این چالش، می‌توانیم مقدار ϵ را درصدی از سوالات که شخص درست پاسخ می‌دهد انتخاب کرد. می‌توان به انواع مسئله وزن نیز اضافه کرد.

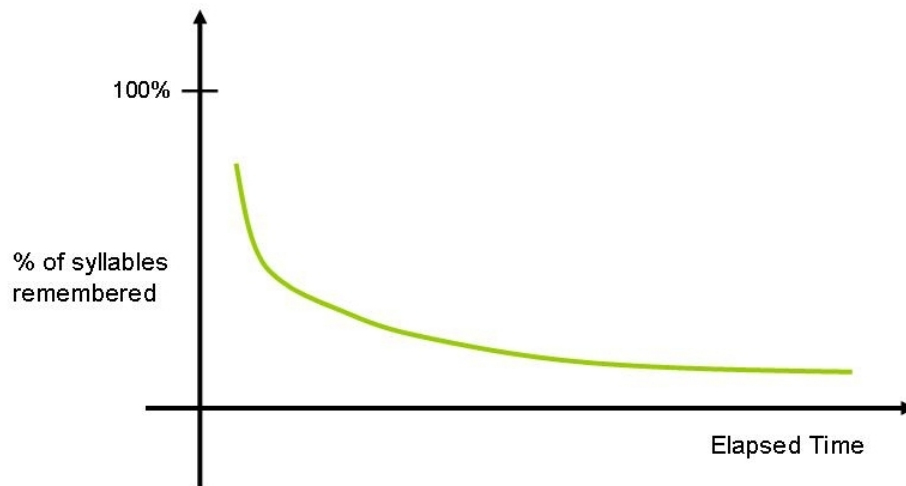
کل فرآیند حدس و جواب نوعی بازی است که یادگیری در آن صورت می‌گیرد. در بازی‌ها، این مسئله حائز اهمیت است که جایزه‌ای که به شخص داده می‌شود، به صورت یک امتیاز است و تنها این امتیاز است که شخص را ترقیب به ادامه‌ی بازی می‌کند. هدف این مدل این است که به ازای هر جواب صحیح، مقداری از یک توکن دیجیتال که مخصوص همین بازی ایجاد شده است به شخص جایزه داده شود. این عمل باعث می‌شود تا علاوه بر مقدار روانی جایزه، وجه مالی آن نیز به انجام بازی ترغیب کند.

همان‌طور که پیش‌تر شرح داده شد، قصد ایجاد نوع یادگیری به صورت ترکیبی از یادگیری همکارانه، شبیه‌سازی و پرسش و پاسخ را داریم. به منظور ایجاد دو المان یادگیری همکارانه و پرسش و پاسخ، می‌توان به این شکل عمل کرد که پس از پاسخ‌دادن به هر سوال، شخص مورد نظر دلایل خود برای انتخاب گزینه را نیز بنویسد. این عمل، برای خود شخص باعث ایجاد reflection می‌شود که طبق مقاله‌ی [] سبب افزایش میزان یادگیری برای خود شخص می‌شود. همچنین، می‌توان دلایل انتخاب اشخاص را به صورت کامنت به هر سوال اضافه کرد تا اشخاص از طریق دیدن دلایل دیگران برای انتخاب گزینه‌ها یادگیری داشته باشند.

۴ مدل‌سازی اشخاص و سناریوها

همان‌طور که پیش‌تر گفته شد، برای هر شخص مقداری تجربه در نظر می‌گیریم که با پارامتر ϵ نمایش داده می‌شود. هدف کلی بازی، افزایش مقدار پارامتر ϵ برای هر شخص است.

مدل کلی یادگیری که می‌توان برای انتخاب سوالات استفاده کرد مدل spaced repetition است. از این مدل به کرات در یادگیری در زمینه‌های مختلف مانند زبان استفاده شده است. بدین منظور، باید در هر سوال تعدادی feature مشخص شود که شخص باید پس از پاسخ‌دادن به سوالات، تاثیر این پارامترها را روی خروجی شناسایی کند و در نتیجه در مراحل بعدی انتخاب صحیحی داشته باشد. در مدل‌های یادگیری ماشین، پس از دیده شدن یک نمونه از داده‌ی یادگیری، انتظار می‌رود که در صورت تکرار آن مدل قادر به پاسخ‌گویی صحیح به سوال باشد. اما در مدل انسانی، شخص دارای پارامتری با عنوان فراموشی است. مدل فراموشی با عنوان Ebbinghaus Model شناخته می‌شود که بیان می‌کند درصد کلمات به یاد مانده بر حسب زمان رابطه‌ای به شکل زیر دارد.



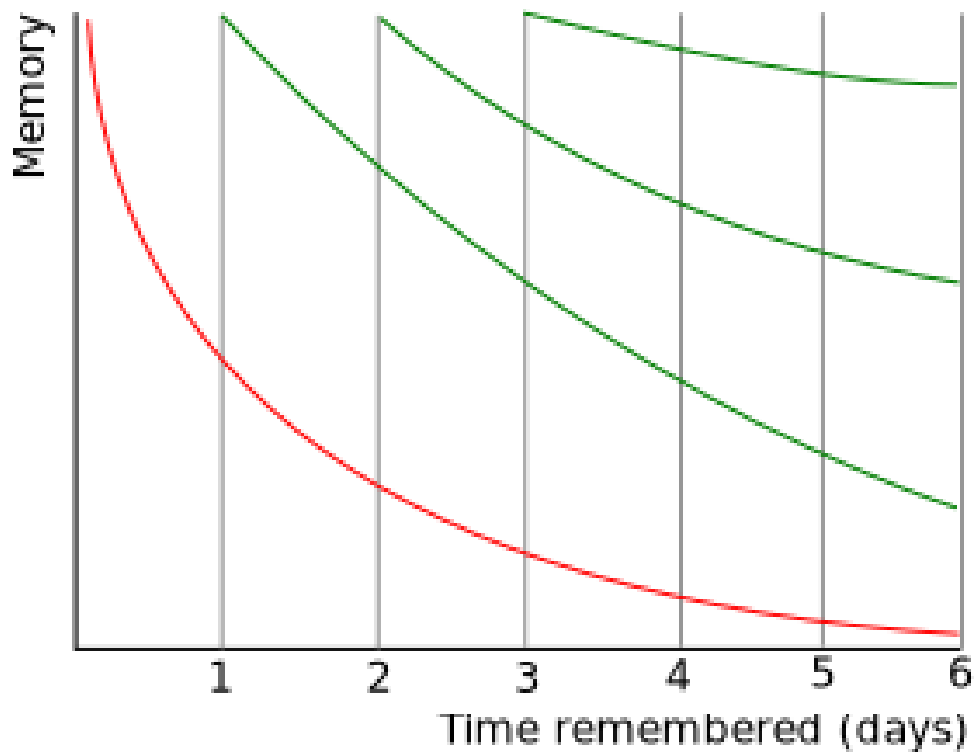
شکل ۱: مدل فراموشی اینگوس

و رابطه‌ی آن به صورت زیر است.

$$b = \frac{100k}{(\log(t))^c + k}$$

در صورت تکرار سوالات به صورت منظم، میزان یادگیری به صورت زیر خواهد بود:

The Forgetting Curve



شکل ۲: فراموشی تکرارشونده

و بعد از تعداد مشخصی تکرار، شیب فراموشی به طرز قابل توجهی کاهش می‌یابد و در این مرحله می‌توان گفت که شخص یک موضوع را یاد گرفته است.
[ویکی‌پدیا و دوئولینگو]

۱.۴ مدل‌های جایزه

۱.۱.۴ جایزه‌ی ثابت

در مدل‌های جایزه‌ای، می‌توان در نظر گرفت که جایزه در زمان دورتر برای شخص نامطلوب‌تر است. فرض کنید در فضای پرسش ترتیبی، ارزش جایزه‌ی سوال‌های بعدی با نرخ γ کاهش می‌یابد. در این صورت، میزان جایزه‌ی دریافت‌شده برای شخص برابر است با:

$$M = \sum_{i=0}^N \epsilon(i) \gamma^i m$$

که در آن $\epsilon(i)$ میزان تخصیص شخص در پاسخ‌گویی در مرحله‌ی i و m میزان جایزه‌ی هر مرحله است. طبق مدل اینگوس، می‌توان بر اساس شرایط شخص انتظاراتی از $\epsilon(i)$ و رفتار کلی آن در طی زمان داشت. مسئله‌ای که در این شرایط پیش می‌آید این است که در صورتی که احتمال پاسخ‌گویی تصادفی درست در هر مرحله برابر p باشد، در صورت پاسخ‌گویی تصادفی در N مرحله، امید ریاضی جایزه‌ی دریافتی به اندازه‌ی زیر است.

$$E(M) = \frac{p(1 - \gamma^N)}{1 - \gamma}$$

بنابراین انتظار می‌رود که تعداد زیادی از اشخاص که دارای تخصیصی کمتر از انتخاب تصادفی هستند، گزینه‌ی تصادفی را انتخاب کنند. به طور کلی می‌توان گفت که این مورد مشکلی ایجاد نمی‌کند و با توجه به هدف شخص از یادگیری تعیین می‌شود، ولی از آنجایی که در این بازی جایزه‌ها به صورت توکن داده می‌شوند، انتخاب تصادفی می‌تواند به ضرر مالی منجر شود. درواقع، تنها زمانی یادگیری تصادفی اتفاق نمی‌افتد که:

$$M > E(M) \rightarrow \sum_{i=0}^N \epsilon(i) \gamma^i m > \frac{p(1 - \gamma^N)}{1 - \gamma}$$

که با توجه به نمودار می‌توان گفت لزوماً اتفاق نمی‌افتد.

۲.۱.۴ محدودیت پاسخ غلط

این مدل در نرم‌افزارهایی مانند Duolingo استفاده می‌شود که هدف آن ترغیب اشخاص به خرید اکانت پولی است. در این مدل، پس از تعدادی جواب غلط شخص تا مدت معینی قادر به پاسخ‌گویی به سوال جدیدی نیست. با این روش می‌توان روند پاسخ‌گویی رندوم را کنترل کرد. ولی همچنان دو مشکل باقی می‌ماند:

۱. سرعت یادگیری برای اشخاص غیرمتخصص کاهش می‌یابد و هدف اصلی که افزایش یادگیری به صورت کلی است محدود می‌شود.

۲. همچنان امید ریاضی پاسخ‌گویی تصادفی از صفر بزرگ‌تر است.

می‌توان شارژشدن دوباره‌ی قابلیت پاسخ‌گویی را به صورت دریافت پول انجام داد که مقدار آن از امید ریاضی پول دریافتی از پاسخ رندوم بیشتر باشد. هرچند این مدل جذابیت چندانی برای مخاطبان ندارد زیرا در این شرایط، یادگیری تا مدت زیادی برای آن‌ها هزینه‌بر است.

۳.۱.۴ مدل شرطبندی

در این مدل هر شخصی به اندازه‌ی مورد نظر خود برای جواب یک سوال پول می‌گذارد. پس از مشخص شدن نتیجه، مجموع پول بین افرادی که پاسخ درستی داده‌اند به صورت مساوی تقسیم می‌شود. با وجودی که این مدل بهترین مدل برای مشارکت در یادگیری است، مسائل مختلفی وجود دارند که باید در آن لحاظ شوند. در ابتدا، امید ریاضی جایزه‌ی هر شخص را با توجه به سطح تخصصش محاسبه می‌کنیم. اگر اشخاص $\{p_1, p_2, \dots, p_n\}$ در بازی شرکت کنند و پولی که به سوال اختصاص می‌دهند $\{m_1, m_2, \dots, m_n\}$ باشد:

$$M = m_1 + m_2 + \dots + m_n$$

$$E(p_i) = \epsilon_i M m_i \sum_{\substack{S \in \{1, \dots, n\} \\ i \notin S}} \frac{\epsilon_S}{m_i + \sum_{\substack{x \in S \\ x \neq i}} m_x}$$

این مدل چند نکته‌ی مهم در پی دارد که باید در پیاده‌سازی به آن‌ها توجه شود.

۱. افراد برای پاسخ دادن به هر سوال باید در دسته‌های مختلفی قرار بگیرند. نوع این دسته‌بندی مهم است. زیرا در صورت یکسان بودن سطح بازیکن‌های سوال، خروجی همه‌ی آن‌ها یکسان خواهد بود ولی برای $\epsilon < 1$ ، لزوماً پول از دست خواهند داد و بازنده خواهند بود.

۲. در صورتی که در دسته‌بندی همه‌ی بازیکن‌های محیط شرکت داشته باشند، سود بازیکن‌های برنده بیشتر و ضرر بازنده‌ها نیز بیشتر خواهد بود.

۳. در صورتی که یک بازیکن با ϵ پایین همیشه در گروه‌هایی با ϵ بزرگ‌تر از خود قرار داشته باشد، احتمالاً به دلیل باخت‌های زیاد انگیزه‌ی یادگیری را از دست می‌دهد.

۴. احتمال باخت مقادیر زیادی از پول برای بازیکن‌ها وجود دارد که انگیزه‌ی یادگیری را کاهش می‌دهد.

برای رفع بخشی از این مشکلات می‌توان راهکارهای متفاوتی را در پیش گرفت. برای مثال، اضافه کردن شرطبندی صفر برای بازیکنان تازه‌وارد یا محدود کردن حداکثر مقدار شرطبندی در بازه‌ی مشخص زمانی با توجه به سطح تخصص بازیکنان.

[بررسی جدول گیم بازی که آیا همه روی صفر شرط می‌بندن یا نه و تعادل شرطبندی کجاست]
به طور کلی، این مدل جمعی از دو مدل قبلی است که همزمان هم شامل برد و هم شامل باخت می‌شود، بازیکنان را به دادن جواب‌های درست تشویق می‌کند و از دادن جواب‌های تصادفی جلوگیری می‌کند. [بررسی بیشتر تاثیر روی یادگیری]

[استفاده از اسمارت کانترکت در سیستم شرطبندی]

[بررسی الگوریتم‌های گروه‌بندی برای بهینه‌سازی هدف افزایش یادگیری]

۲.۴ بخش‌های اضافی بازی

از آنجایی که هدف بازی یادگیری نحوه‌ی سرمایه‌گذاری است، یک محیط جهت معامله در بازی نیز فراهم می‌شود. بازیکنان می‌توانند در این محیط توکن‌های خود را خرید و فروش کنند و تجربیات کسب‌شده‌ی معاملات خود را در عمل بررسی کنند. این خرید و فروش می‌تواند منبع اصلی کسب توکن برای شروع به یادگیری و سرمایه‌گذاری روی سوالات باشد و خرید و فروش با پول واقعی یا مکانیزمی دیگر انجام شود.

۳.۴ انتخاب فریم سوالات

انتخاب بازه‌ی زمانی و داده‌های مورد نمایش در صورت سوال می‌تواند به صورت تصادفی صورت گیرد. اما در صورت انتخاب هدفمند احتمالاً نتیجه‌ی بهتری حاصل می‌شود. در روش هدفمند، ابتدایستی از featureهای مورد نظر برای یادگیری انتخاب می‌شود و سپس تعدادی الگوریتم برای پیدا کردن این فیچرها در داده‌های مختلف طراحی می‌شوند. سپس سوالات با استفاده از این الگوریتم‌ها انتخاب می‌شوند و به اشخاص نمایش داده می‌شوند.

۵ شبیه‌سازی

پس از پیاده‌سازی سیستم، نیاز به بررسی عملکرد آن در رسیدن به اهداف ذکر شده داریم. بدین منظور، دو راهکار قابل انجام است. راهکار اول استفاده از آزمایش‌های انسانی است، که هزینه‌ی بالا و سختی زیادی دارند و در صورت نیاز به تغییر عملکرد و پارامترهای سیستم طبق مشاهدات، هزینه‌ها بیشتر هم خواهند شد. راهکار دیگر شبیه‌سازی agentهای درون سیستم، بهینه‌سازی طبق رفتار آن‌ها و سپس در صورت نیاز شروع به آزمایش انسانی است. برای شبیه‌سازی agentها چند راهکار مختلف قابل اتخاذ است.

۱.۵ ایجنت رندوم

این ایجنت‌ها، بدون داشتن هیچ دانش قبلی و به صورت تصادفی به سوالات پاسخ می‌دهند. با این که به طور کلی ایجنت‌های از این دست فاقد ارزش در بحث یادگیری هستند. اما از آنجایی که در همه‌ی سیستم‌ها افرادی به این شکل وجود دارند، در هر نوع مدل‌سازی افزودن این مدل افراد می‌تواند به شبیه‌سازی بهتر سیستم کمک کند. فقط باید تعریف دقیقی برای رندوم بودن یا نبودن ایجنت تعریف شود و رفتارهای آن مشخص شود تا در سیستم انسانی نیز بتوان از این مدل استفاده کرد.

۲.۵ ایجنت مدل‌شده

راه دیگر استفاده از مدل‌های روان‌شناسی برای شبیه‌سازی افراد هوشمند در سیستم است. برای مثال استفاده از مدل فراموشی اینگوس با پارامترهای تصادفی متفاوت. مشکلی که در این سیستم وجود دارد این است که رفتار افراد از پیش از شبیه‌سازی مشخص است، و شبیه‌سازی ممکن است کمکی به درک بهتر رفتار اشخاص در سیستم نکند. همچنین ممکن است مدل‌های روان‌شناسی کاملاً درست نباشند و سیستم نتواند رفتار انسانی را به درستی نشان دهد. اما از مزیت‌های این مدل این است که در صورت وجود مدل دقیق، پیاده‌سازی آن به نسبت راحت است.

۳.۵ یادگیری

راهکار آخر استفاده از مدل‌های یادگیری ماشین است. برای این مسئله، می‌توان از یادگیری تقویتی استفاده کرد و از مدل‌سازی اضافی پرهیز کرد. یکی از مسائل این گونه مدل‌ها این است که رفتار آن‌ها با رفتار انسان کاملاً منطبق نیست، و بنابراین باید تعدادی پارامتر به مسئله اضافه کرد تا بتوانند شبیه‌تر به انسان رفتار کنند. برای مثال، باید پارامتری با عنوان فراموشی به ایجنت RL اضافه کرد تا این رفتار انسانی را مدل کند. از مزیت‌های دیگر مدل تقویتی می‌تواند این باشد که علاوه بر مدل‌سازی سیستم، در صورت یادگیری درست ایجنت‌ها خود آن‌ها می‌توانند برای پیش‌بینی رفتار بازار و به عنوان سیستم‌های خبره استفاده شوند.

یکی از نکاتی که باید در هر دو مدل آخر رعایت کرد، اضافه کردن پارامترهایی برای شبیه‌سازی رفتار اجتماعی ایجنت‌ها است. همان‌طور که پیش‌تر ذکر شد، یکی از اهداف مدل یادگیری استفاده از خودآموزی

و یادگیری اجتماعی است. از آنجایی که شبیه‌سازی مدل خودآموزی تا حدی غیرممکن است، و شاید تنها بتوان در پارامترهای یادگیری آن را لحاظ کرد، باید روی مکانیزم‌های یادگیری اجتماعی بیشتر بررسی صورت گیرد.