



کاربرد داده کاوی در حوزه مالی

شرکت عصر امین کارآفرین

فهیمة روشن، فاطمة پازوکی

فهرست

۳.....	چکیده
۳.....	مقدمه
۳.....	روش تحقیق
۴.....	تعریف داده کاوی
۴.....	کاربرد علم داده کاوی در بازارهای مالی
۸.....	روش‌های داده کاوی
۱۱.....	مراحل داده کاوی
۱۲.....	نمونه‌ای از شرکت‌های فین‌تکی
۱۳.....	نتیجه‌گیری
۱۴.....	منابع

چکیده

امروزه نقش و اهمیت داده‌ها به عنوان منبعی ارزشمند و عاملی کلیدی برای رشد کسب و کارهای مختلف و ایجاد مزیت رقابتی در بازارهای متعدد برکسی پوشیده نیست. رشد و توسعه روزافزون مجموعه داده‌ها منجر به بزرگتر و پیچیده تر شدن آنها شده بطوری که پایگاه‌های داده‌ای سنتی قادر به پوشش آنها نبوده و ابزارهای سنتی نمی‌توانند آنها را مورد تجزیه و تحلیل قرار دهند.

مقاله حاضر به بررسی اهمیت تحلیل داده در صنعت مالی و تأثیر آن بر تصمیم‌گیری‌های سرمایه‌گذاری می‌پردازد. این تحقیق با تمرکز بر روش‌ها، الگوریتم‌ها و مدل‌های تحلیل داده در حوزه مالی، نقش بیانگر آن در بهبود دقت پیش‌بینی‌ها و کاهش ریسک‌های مرتبط با سرمایه‌گذاری، را برجسته می‌کند. همچنین تعدادی از شرکت‌های مالی که به صورت جدی در این زمینه فعالیت دارند بررسی گردیده است.

مقدمه

تحلیل داده در صنعت مالی نقش بسیار حیاتی دارد. این امر مرتبط با دو مسئله اساسی هست: اطلاعات فراوان و تصمیم‌گیری دقیق. اولاً، بازارهای مالی اطلاعات بی‌شماری تولید می‌کنند که شامل داده‌های قیمت‌ها، اخبار، اطلاعات مالی شرکت‌ها، و غیره هستند. تحلیل این داده‌ها به سرمایه‌گذاران امکان می‌دهد تا الگوهای بازار و روندهای مهم را شناسایی کنند. دوماً، تحلیل داده با استفاده از الگوریتم‌ها و مدل‌های پیشرفته می‌تواند به سرمایه‌گذاران کمک کند تا تصمیم‌های موثرتری برای سرمایه‌گذاری بگیرند. این مدل‌ها می‌توانند الگوهای پیشین را مورد بررسی قرار بدهند و با استفاده از آنها به پیش‌بینی تغییرات بازار و ریسک‌ها کمک کنند. این تکنیک‌ها می‌توانند به دستیابی به استراتژی‌های مناسب برای سرمایه‌گذاری کمک کنند، که در نهایت منجر به بهبود عملکرد و دستیابی به اهداف سرمایه‌گذاری می‌شود.

روش تحقیق

این مطالعه بر اساس داده‌های ثانویه است و مطالب از مقالات، کتاب‌ها، مجلات، روزنامه‌ها و وب‌سایت‌های مربوطه جمع‌آوری شده است.

تعریف داده کاوی

- داده کاوی استخراج نیمه اتوماتیک الگوها، تغییرات، وابستگی ها، ناهنجاری ها و دیگر ساختارهای معنی دار آماری از پایگاه های بزرگ داده می باشد.
- داده کاوی علم استخراج اطلاعات مفید از پایگاه های داده یا مجموعه داده ای می باشد
- داده کاوی عبارت است از فرایند استخراج اطلاعات معتبر، از پیش ناشناخته، قابل فهم و قابل اعتماد از پایگاه داده های بزرگ و استفاده از آن در تصمیم گیری در فعالیتهای تجاری مهم
- اصطلاح داده کاوی به فرایند نیم خودکار تجزیه و تحلیل پایگاه داده های بزرگ به منظور یافتن الگوهای مفید اطلاق می شود
- داده کاوی یعنی جستجو در یک پایگاه داده ها برای یافتن الگوهایی میان داده ها
- داده کاوی استخراج اطلاعات مفهومی، ناشناخته و به صورت بالقوه مفید از پایگاه داده می باشد

کاربرد علم داده کاوی در بازارهای مالی

علم داده در بازارهای مالی و سرمایه گذاری بسیار کاربردی و مورد توجه قرار دارد. با تحلیل دقیق داده ها و کنترل اطلاعات قبلی، می توان به تحلیل و شناختی از داده های آینده دست یافت. حجم بزرگی از داده های مالی هر ثانیه تولید می شود و پایش و تحلیل این داده ها فرصت های مناسبی برای کسب سود در بازار مالی و سرمایه گذاری ایجاد می کند. با استفاده از داده کاوی بر روی این داده ها، می توان اطلاعات مفیدی را استخراج کرده و برای کنترل نوسانات بازار و تصمیم گیری درست استفاده کرد. علم داده در بازارهای مالی و سرمایه گذاری به عنوان یک مرحله ابتدایی تصمیم گیری مهم قلمداد می شود و با ورود داده های درست، نتایج دقیقی را می توان از آن دریافت کرد.

تحلیل داده در حوزه مالی به استفاده از داده های مالی و اقتصادی به منظور استخراج اطلاعات، الگوها، روندهای مفید برای تصمیم گیری در بازارهای مالی می پردازد. این فرایند شامل استفاده از روش ها و الگوریتم های مختلف برای تجزیه و تحلیل داده ها به منظور پیش بینی قیمت ها، ارزیابی ریسک، شناسایی الگوهای بازار، و ارائه راهکارهای بهبود عملکرد در سرمایه گذاری ها می باشد.

این امر به سرمایه گذاران و صنعت های مالی امکان می دهد تا تصمیمات بر اساس اطلاعات دقیق تر و پیش بینی های بهتری بگیرند.

اکنون از قیمت گذاری خودکار گرفته تا بانکداری آنلاین شخصی، همه چیز ارائه می گردد. تعدادی استفاده برتر علم داده را در صنعت مالی بررسی می کنیم:

۱- بینش بازار سهام در زمان حال

نقش داده‌ها در بازار سهام همیشه مهم بوده است، از نظر تاریخی، پیگیری سهام و خرید و فروش آن به معنای تجزیه و تحلیل داده‌های گذشته با دست بود. این روش به سرمایه‌گذاران اجازه می‌داد تا بهترین تصمیمات ممکن را بگیرند، اما یک رویکرد ناقص بود. این روش نوسانات بازار را در نظر نمی‌گرفت، به این معنی که معامله‌گران فقط می‌توانند از داده‌هایی که به صورت دستی ردیابی و اندازه‌گیری شده‌اند، همراه با شهود شخصی استفاده کنند. جای تعجب نیست که تصمیمات بد سرمایه‌گذاری با استفاده از داده‌های منسوخ، چندان غیرمعمول نبود.

امروزه، با استفاده از پیشرفت‌های فناوری، دانشمندان داده‌های مالی (با تمام اهداف عملی) این تأخیر داده‌ها را ریشه‌کن کرده و جریان دائمی بینش در زمان حال را برای ما فراهم کرده‌اند. با استفاده از خطوط انتقال داده پویا، معامله‌گران می‌توانند به اطلاعات بازار سهام در زمان وقوع دسترسی پیدا کنند. با ردیابی معاملات در زمان حال، آن‌ها می‌توانند تصمیمات بسیار هوشمندانه‌تری در مورد خرید و فروش سهام بگیرند که حاشیه خطا را بسیار کاهش می‌دهد.

۲- معاملات الگوریتمی

هدف از معاملات بازار سهام خرید سهام با قیمت پایین است. این مورد شامل استفاده از روند گذشته و حال بازار است تا بدانید کدام سهام احتمال افزایش یا کاهش قیمت دارد. برای به حداکثر رساندن سود، معامله‌گران بورس باید سریع وارد شوند و قبل از رقبا سهام خود را خریداری کرده و بفروشند. این کار قبلاً به صورت دستی انجام می‌شد؛ اما با ورود کلان داده و بینش‌های زمان حال، این روند تغییر شکل داده است. معاملات الگوریتمی وارد بازی شدند. دانشمندان داده‌های مالی با استفاده از الگوریتم‌های یادگیری ماشین (Machine Learning) و داده‌های موجود، یک نوع معامله کاملاً جدید ایجاد کرده‌اند: معامله با تناوب بالا (HFQ) از آنجا که روند کار کاملاً خودکار است، خرید و فروش می‌تواند با سرعت بسیار بالا اتفاق بیفتد.

۳- مدیریت ریسک خودکار

مدیریت ریسک مالی به معنی محافظت از سازمان‌ها در برابر تهدیدات احتمالی است. تهدیدات می‌تواند دامنه وسیعی داشته باشد و شامل مواردی مانند ریسک اعتباری (به عنوان مثال؛ آیا این مشتری پرداخت کارت خود را به طور پیش‌فرض انجام می‌دهد؟) و ریسک بازار (به عنوان مثال آیا حباب مسکن می‌ترکد؟) می‌شود. انواع دیگر شامل ریسک تورم، ریسک قانونی و غیره است. اساساً، هر چیزی که ممکن است بر عملکرد یا سود موسسه مالی تأثیر منفی بگذارد، می‌تواند یک ریسک تلقی شود.

مدیریت ریسک شامل سه وظیفه است: تشخیص ریسک‌ها، نظارت بر آن‌ها، و اولویت‌بندی اینکه با کدام ریسک‌ها باید با فوریت بیشتری برخورد گردد. این امر ممکن است ساده به نظر برسد، اما وقتی همه عوامل ریسک و نحوه تلاقی آن‌ها را در نظر بگیرید، به سرعت پیچیده می‌شود. دانشمندان داده با خودکارسازی شناسایی، نظارت و اولویت‌بندی ریسک، الگوریتم‌های یادگیری ماشین دامنه خطاهای انسانی را به حداقل می‌رسانند. آن‌ها همچنین طیف گسترده‌ای از منابع مختلف داده (از داده‌های مالی گرفته تا داده‌های بازار و رسانه‌های اجتماعی مشتری) را در نظر می‌گیرند و تأثیر این منابع مختلف بر یکدیگر را اندازه‌گیری می‌کنند. برای مثال، شرکت‌های کارت اعتباری با استفاده از نرم‌افزار خودکار مدیریت ریسک اکنون می‌توانند قابلیت اطمینان مشتری بالقوه را به طور دقیق تعیین کنند، حتی اگر زمینه مالی جامع مشتری را نداشته باشند. یک مزیت این الگوریتم‌ها این است که هرچه رشد می‌کنند، بهبود می‌یابند. مدیریت ریسک مبتنی بر هوش مصنوعی Artificial Intelligence و پذیره‌نویسی هوشمند می‌تواند ارتباطاتی ایجاد کند که بشر به تنهایی هرگز آن را پیدا نمی‌کند. این قدرت یادگیری ماشین است. در حالی که این رویکردها در صنعت مالی نسبتاً جدید هستند، اما پتانسیل آن‌ها برای آینده بسیار زیاد است.

۴- کشف تقلب و جعل

کلاهبرداری مالی اشکال مختلفی دارد؛ از جمله جعل کارت‌های اعتباری، ادعاهای بیمه‌ای و جرایم سازمان‌یافته. جلوگیری از وقوع کلاهبرداری برای هر موسسه مالی حیاتی است. این مورد فقط شامل به حداقل رساندن خسارات مالی نیست، بلکه در مورد اعتماد است. بانک‌ها مسئولیت اطمینان از امنیت مشتریان خود را دارند. با استفاده از داده‌کاوی، دانشمندان داده می‌توانند ناهنجاری‌ها یا الگوهای غیرمعمول را در صورت بروز تشخیص دهند. سپس الگوریتم‌هایی که به طور خاص طراحی شده‌اند، موسسه را از رفتار ناهنجار آگاه و به طور خودکار فعالیت مشکوک را مسدود می‌کنند. بارزترین نمونه آن کلاهبرداری در کارت اعتباری است. به عنوان مثال، اگر کارت شما در یک مکان غیرمعمول مورد استفاده قرار گیرد یا برداشت به شکلی انجام شود که معمولاً توسط کلاهبرداران انجام می‌گردد، شرکت کارت اعتباری می‌تواند کارت را مسدود کرده و به شما اطلاع دهد که مشکلی پیش آمده است. یادگیری ماشین همچنین می‌تواند الگوهای گسترده‌تری از رفتار ناهنجار را تشخیص دهد، به عنوان مثال سازمان‌های مختلفی که به طور هم‌زمان هک می‌شوند. این مورد می‌تواند به بانک‌ها کمک کند تا حملات سایبری و جرایم سازمان‌یافته را شناسایی کنند و میلیون‌ها نفر را نجات دهند.

۵- تجزیه و تحلیل مصرف‌کننده

برای هر بانک یا ارائه‌دهنده خدمات مالی، درک رفتار مشتری جهت تصمیم‌گیری صحیح حیاتی است. با توجه به منابع مختلف داده و استفاده از اطلاعات جمعیتی مانند سن و موقعیت جغرافیایی، بانک‌ها، شرکت‌های بیمه،

صندوق‌های بازنشستگی و شرکت‌های کارت اعتباری می‌توانند بینش بسیار دقیقی داشته باشند. با استفاده از این بینش‌ها، آن‌ها می‌توانند بازاریابی مستقیم و رویکرد مدیریت ارتباط با مشتری را بر این اساس مطابقت دهند. این ممکن است شامل استفاده از داده‌ها برای فروش محصولات خاص یا بهبود خدمات مشتری باشد.

۶- خدمات شخصی‌سازی شده

اولین مورد شخصی‌سازی نسخه‌ای است. این مورد از داده‌ها و ترجیحات برگزیده مشتری در گذشته را برای پیش‌بینی نیازهای آن‌ها استفاده می‌کند. به طور کلی توسط الگوریتم‌هایی هدایت می‌شود که به تعاملات مشتری پاسخ می‌دهند.

نوع دوم شخصی‌سازی در زمان واقعی است. این مورد به اطلاعات قبلی و فعلی متکی است تا تجربه مشتری را در هنگام وقوع (برای مثال اگر در هنگام انجام معامله آنلاین، کالا یا خدماتی به شما پیشنهاد می‌شود) تنظیم کنید. نوع نهمی، شخصی‌سازی یادگیری ماشین است. اگرچه این یک مفهوم نسبتاً جدید می‌باشد، اما در حال حاضر دارای پتانسیل جالب توجهی است. یک مثال عالی، نرم‌افزار wallet.AI است که از مشخصات مالی و سابقه معاملات کاربران استفاده کرده و به عنوان مشاور شخصی در هزینه‌های روزانه عمل می‌کند.

۷- قیمت‌گذاری و بهینه‌سازی درآمد

بهینه‌سازی قیمت‌گذاری توانایی شکل دادن قیمت‌گذاری بر اساس زمینه‌ای است که مشتری با آن روبرو می‌شود. بیشتر بانک‌ها و ارائه‌دهندگان بیمه تیم‌های فروش بزرگی دارند و شبکه‌های پیچیده‌ای از محصولات و خدمات مختلف را ارائه می‌دهند. دانشمندان داده‌های مالی با استفاده از داده‌های مختلف از منابعی مانند نظرسنجی‌ها، قیمت‌گذاری محصولات گذشته و تاریخچه فروش، می‌توانند به سودآوری و صرفه‌جویی در وقت و نیرو برای این تیم‌های فروش کمک کنند.

۸- توسعه محصول

یکی از موارد استفاده علم داده در صنعت مالی، ن (fintech) فناوری مالی است. این حوزه نوپا در صنعت در سال‌های اخیر رشد قابل توجهی داشته است، شرکت‌های فناوری مالی با سرعت بسیار بالاتری از آنچه سازمان‌های جهانی قادر به مدیریت آن هستند، نوآوری‌های هیجان‌انگیزی را ارائه می‌دهند. در حالی که بسیاری از ارائه‌دهندگان فناوری مالی بانک‌های دیجیتال را راه‌اندازی کرده‌اند، برخی دیگر قبل از فروش این موارد، روی حوزه خاصی از فناوری تمرکز می‌کنند. بلاکچین و ارز رمزنگاری شده، سیستم‌عامل‌های پرداخت تلفن همراه، برنامه‌های معامله مبتنی

بر تجزیه و تحلیل، نرم افزار و امدهی و محصولات بیمه‌ای مبتنی بر هوش مصنوعی تنها چند نمونه از فناوری مالی هستند که توسط علم داده هدایت می‌شوند.

۹- مدیریت عمومی داده‌ها

همانطور که گفته شد، موسسات مالی به حجم عظیمی از داده‌ها دسترسی دارند. این داده‌ها از منابع گسترده‌ای به دست می‌آیند: معاملات تلفن همراه، داده‌های رسانه‌های اجتماعی، معاملات نقدی، گزارش‌های بازار و... علاوه بر غول‌های رسانه‌های اجتماعی، بخش مالی بیش از هر صنعت دیگر به داده‌های ما دسترسی دارد. مدیریت داده در امور مالی به تیم‌هایی متشکل از متخصصان داده احتیاج دارد که بتوانند داده‌ها را استخراج و پیچیدگی‌های صنعت را درک کنند

مهمترین الگوها و روش‌های استفاده شده تحلیل داده برای پیش‌بینی در حوزه مالی عبارتند از:

الگوهای فنی: استفاده از الگوهای نمودارهای مالی، میانگین‌ها، شاخص‌های فنی مثل MACD یا RSI برای پیش‌بینی رفتار قیمت‌ها.

تحلیل پایه‌ای: بررسی عوامل اقتصادی، اطلاعات شرکتی، و اخبار مرتبط برای پیش‌بینی تغییرات در قیمت‌ها. مدل‌های پیشرفته یادگیری ماشین: استفاده از رگرسیون، شبکه‌های عصبی و سایر مدل‌های یادگیری ماشین برای پیش‌بینی دقیق‌تر قیمت‌ها. تحلیل دقیق داده‌ها به همراه استفاده از مدل‌های پیشرفته یادگیری ماشین می‌تواند به سرمایه‌گذاران کمک کند تا الگوها و روندهای مخفی در داده‌ها را شناسایی کنند و به‌طور دقیق‌تر قیمت‌ها را پیش‌بینی کنند.

در ادامه به بررسی چند نمونه مدل‌های پیشرفته یادگیری ماشین می‌پردازیم:

الگوریتم‌های یادگیری ماشین یا "Machine Learning" دسته‌ای از الگوریتم‌ها و تکنیک‌های محاسباتی هستند که به کامپیوترها امکان می‌دهند از داده‌ها یاد بگیرند و الگوها یا روابطی را در داده‌ها شناسایی کنند بدون نیاز به برنامه‌ریزی صریح.

این الگوریتم‌ها به دو دسته اصلی تقسیم می‌شوند:

• یادگیری نظارت شده: (Supervised Learning)

در این روش، مدل‌ها با استفاده از داده‌های آموزشی که دارای ورودی و خروجی مشخص هستند، یاد می‌گیرند. به عبارت دیگر، آن‌ها از داده‌های ورودی و مرتبط آن‌ها با خروجی یاد می‌گیرند تا بتوانند پیش‌بینی‌هایی برای داده‌های جدید ارائه کنند.

• یادگیری بدون نظارت: (Unsupervised Learning)

در این روش، مدل‌ها از داده‌هایی که بدون برچسب یا خروجی مشخص هستند یاد می‌گیرند. آن‌ها سعی می‌کنند الگوها و ساختارهای مخفی در داده‌ها را شناسایی کرده و گروه‌بندی یا تحلیل دقیق‌تری از داده‌ها ارائه دهند. این تکنیک‌ها برای پردازش داده‌های مالی، شناسایی الگوها در بازارهای مالی و پیش‌بینی تغییرات و روندهای مختلف مورد استفاده قرار می‌گیرند و به سرمایه‌گذاران کمک می‌کنند تا تصمیمات بهتری بگیرند

شبکه‌های عصبی: (Neural Networks)

این مدل‌ها براساس ساختار مغز انسان عمل می‌کنند در حوزه تحلیل داده، شبکه‌های عصبی برای تحلیل دقیق داده‌های مالی و پیش‌بینی روندهای بازار استفاده می‌شوند. این شبکه‌ها به واسطه قابلیتشان در تشخیص الگوها و ارتباطات پیچیده در داده‌ها، می‌توانند الگوهای مخفی و روندهای پیچیده را در بازارهای مالی شناسایی کنند. همچنین، این شبکه‌ها می‌توانند از داده‌های مالی گذشته و فعلی بهره‌برده و الگوهای تکراری یا پیش‌بینی‌های مبتنی بر ریسک و عملکرد را بررسی کنند. از طریق آموزش این شبکه‌ها با داده‌های مالی، می‌توان به پیش‌بینی قیمت‌ها، رفتار بازار، و تصمیمات سرمایه‌گذاری کمک کرد

طبقه‌بندی: (Classification)

طبقه‌بندی یکی از روش‌های مهم تحلیل داده است که در حوزه فین‌تک کاربردهای زیادی دارد. این روش به تقسیم داده‌ها به گروه‌های مختلف یا دسته‌هایی با ویژگی‌های مشترک کمک می‌کند.

در مالی، طبقه‌بندی برای موارد مختلف استفاده می‌شود:

پیش‌بینی رفتار بازار: این روش می‌تواند سهام یا سایر اوراق بهادار را بر اساس ویژگی‌های مشترک در گروه‌های مختلف طبقه‌بندی کند.

- تشخیص سرمایه‌گذاری‌های خطرناک: طبقه‌بندی می‌تواند به شناسایی سرمایه‌گذاری‌هایی که ممکن است ریسک بالا داشته باشند کمک کند.

- تشخیص تقلب: در مالی، این روش برای تشخیص معاملات تقلبی یا فعالیت‌های غیرقانونی استفاده می‌شود.

- تحلیل اعتباری: برای ارزیابی ریسک و اعتبار وام‌دهی، طبقه‌بندی بر اساس ویژگی‌های مختلف انجام می‌شود.

رگرسیون (Regression) : رگرسیون یکی از تکنیک‌های مهم تحلیل داده است که برای پیش‌بینی مقادیر عددی براساس روابط میان متغیرها استفاده می‌شود. در حوزه مالی، رگرسیون به طور گسترده برای پیش‌بینی و تحلیل داده‌های مالی مورد استفاده قرار می‌گیرد.

رگرسیون در تحلیل داده با بررسی رابطه بین یک متغیر وابسته (متغیر واقعی که می‌خواهیم پیش‌بینی کنیم، مانند قیمت سهام) و یک یا چند متغیر مستقل (متغیرهایی که فرض می‌شود تأثیرگذار بر متغیر وابسته هستند، مانند فاکتورهای اقتصادی) انجام می‌شود.

برای این کار، رگرسیون از داده‌های موجود استفاده می‌کند و سعی می‌کند رابطه‌ای ریاضی بین متغیرهای وابسته و مستقل را مدل کند. این مدل رابطه‌ای است که بعداً می‌تواند برای پیش‌بینی مقادیر جدید متغیر وابسته بکار رود. در مالی، مثلاً اگر بخواهیم قیمت یک سهم را بر اساس چند عامل مانند درآمد شرکت، اخبار صنعت، و روند بازار پیش‌بینی کنیم، از رگرسیون برای بررسی و تعیین ارتباطات ریاضی بین این متغیرها استفاده می‌شود. سپس، این رابطه می‌تواند برای پیش‌بینی قیمت‌های آتی سهم به کار رود.

تحلیل سرمایه‌گذاری (Portfolio Analysis): تحلیل پرتفوی یا Portfolio Analysis یک روش است که در آن مجموعه‌ای از دارایی‌ها یا سرمایه‌های سرمایه‌گذاری تحت مجموعه‌ای از معیارها و شرایط تحلیل می‌شود تا بهینه‌سازی ترکیب این دارایی‌ها و توزیع سرمایه را فراهم کند.

در تحلیل پرتفوی، مهمترین اهداف عبارتند از:

- تنوع در سرمایه‌گذاری: این روش برای ایجاد ترکیبی از دارایی‌ها با ریسک و بازده مختلف بر اساس اهداف سرمایه‌گذاری استفاده می‌شود.
- بهره‌وری مالی: هدف از تحلیل پرتفوی بهینه‌سازی ترکیب دارایی‌ها به نحوی است که ریسک کلی کاهش یابد و بازده سرمایه‌گذاری افزایش یابد.
- مدیریت ریسک: از این روش برای کاهش ریسک‌های مرتبط با سرمایه‌گذاری در بازارهای مالی استفاده می‌شود.

تحلیل پرتفوی یا Portfolio Analysis با مراحل زیر انجام می‌شود:

- انتخاب دارایی‌ها: در این مرحله، دارایی‌های مختلفی که سرمایه‌گذاران ممکن است در نظر داشته باشند، انتخاب می‌شوند. این دارایی‌ها ممکن است شامل سهام، اوراق بهادار، ارزها یا هر دارایی دیگری باشند.
- ارزیابی ریسک و بازده: سپس، با استفاده از مدل‌ها و روش‌های مختلف، ریسک و بازده پتانسیلی هر دارایی مورد بررسی قرار می‌گیرد.
- ترکیب دارایی‌ها: بعد از انجام ارزیابی، دارایی‌ها به نحوی ترکیب می‌شوند که بازده مطلوب و ریسک قابل قبول را برای سرمایه‌گذار فراهم کنند. این مرحله شامل تعیین وزن دارایی‌ها در پرتفوی است.

- ارزیابی و بهینه‌سازی: سپس پرتفوی ایجاد شده ارزیابی می‌شود و در صورت نیاز تغییراتی برای بهبود ریسک و بازده ممکن است اعمال شود.
- پیگیری و تنظیم: به مرور زمان، پرتفوی تنظیم می‌شود و با تغییر شرایط بازار و اهداف سرمایه‌گذاری، ممکن است نیاز به تغییرات داشته باشد.

دانشمندان داده می‌توانند به چندین روش مختلف در این زمینه کمک کنند:

- توسعه مدل‌های بهتر برای پیش‌بینی قیمت دارایی‌ها و دیگر شاخص‌های بازار؛
- شناسایی عوامل خطر جدید که ممکن است در داده‌های سنتی ظاهر نشوند؛
- اتوماسیون اجرای معاملات بر اساس پیش‌بینی‌های مدل؛
- ایجاد هشدارهای فوری هنگام وقوع حرکات غیرمنتظره در بازار.

این مسائل نشان‌دهنده اهمیت بی‌پایان علم داده در بهبود روش‌های مدیریت سرمایه‌گذاری و ایجاد تصمیمات بهینه در صنعت مالی هستند. این فرآیند تحلیل دقیق پرتفوی و بهره‌وری مالی سرمایه‌گذاران را افزایش می‌دهد و به آن‌ها کمک می‌کند تا ترکیبی از دارایی‌ها را انتخاب کنند که با هدف‌ها و ریسک‌های آن‌ها سازگاری داشته باشد.

مراحل (فرآیند) داده‌کاوی:

- ۱- شناسایی هدف: در این مرحله مشخص می‌شود کاربر چه چیزی را می‌خواهد و تا چه سطحی از اطلاعات رادر نظر دارد که از پایگاه داده، اخذ نماید
- ۲- انتخاب داده‌ها: در این مرحله باید داده‌ها بر مبنای معیارهای مشخص انتخاب گردند.
- ۳- آماده‌سازی داده‌ها: شکل قابل استفاده داده و شناسایی متغیرهای زائد و وظیفه این مرحله از فرآیند خواهد بود
- ۴- ارزیابی داده‌ها: چارچوب کلی این مرحله، معیارهایی از قبیل نوع توزیع داده‌ها، ویژگی‌ها و ساختار پایگاه داده و شرایط کلی داده‌ها و... می‌باشد
- ۵- قالب‌بندی پاسخ: خروجی این بخش، ارائه فرمت به شکل تصویر، نمودار، شبکه عصبی و... است
- ۶- انتخاب ابزار: در این مرحله ابزارهای مناسب برای داده‌کاوی انتخاب می‌گردد
- ۷- الگوسازی: فرآیند داده‌کاوی به صورت اصلی از این مرحله آغاز می‌گردد که شامل جستجوی الگوها در مجموعه داده، طبقه‌بندی و ارزشیابی داده‌ها و... می‌باشد.
- ۸- اعتبارسازی یافته‌ها: این مرحله، شامل آزمون کردن الگوها است
- ۹- ارزیابی نتایج: نتیجه این بخش، گزارش نهایی برای کاربر است.

- ۱۰ - استفاده از نتایج: هدف اصلی داده کاوی استفاده از نتایج کشف شده برای تصمیم گیری، سیاستگذاری و پیشبینی به منظور ایجاد یک موقعیت بهتر و جدید می باشد.
- موارد زیر نمونه هایی از شرکت هایی هستند که در حوزه فین تک فعالیت دارند و از تحلیل داده به صورت جدی استفاده می کنند:
- Revolut:** این شرکت فین تک به عنوان یک اپلیکیشن بانکی دیجیتال، از تحلیل داده برای ارائه خدمات مالی به مشتریان استفاده می کند.
- Robinhood:** این پلتفرم معاملاتی از تحلیل داده برای فهم بهتر رفتار مشتریان و شناسایی روندها در بازارهای مالی استفاده می کند.
- Plaid:** به عنوان یک شرکت فین تک، Plaid به تحلیل داده برای ارائه خدمات اتصال به حساب های بانکی مشتریان استفاده می کند.
- Square:** این شرکت فین تک از تحلیل داده برای فراهم کردن خدمات پرداخت و مدیریت مالی برای کسب و کارها استفاده می کند.
- Betterment:** این شرکت فین تک، به عنوان یک سرویس مشاوره سرمایه گذاری، از تحلیل داده برای پیشنهاد سرمایه گذاری های بهینه برای مشتریان استفاده می کند.
- Adyen:** این شرکت در حوزه پرداخت ها فعالیت دارد و از تحلیل داده برای بهبود خدمات پرداخت و پشتیبانی مشتریان استفاده می کند.
- Ant Financial:** این شرکت مالی چینی، از تحلیل داده برای ارائه خدمات بانکی دیجیتال و پرداخت استفاده می کند.
- Wealthfront:** به عنوان یک شرکت فین تک مشاوره سرمایه گذاری، از تحلیل داده برای پیشنهاد سرمایه گذاری های مبتنی بر هدف استفاده می کند.
- Klarna:** این شرکت فین تک از تحلیل داده برای ارائه خدمات پرداخت و خرید آنلاین استفاده می کند.
- Stripe:** این پلتفرم پرداخت از تحلیل داده برای بهبود فرآیند پرداخت و ارائه خدمات پرداخت به کسب و کارها استفاده می کند.
- Chime:** این شرکت از تحلیل داده برای ارائه خدمات بانکی دیجیتال و مدیریت مالی فردی استفاده می کند.
- SoFi:** به عنوان یک پلتفرم مالی چندگانه، SoFi از تحلیل داده برای ارائه خدمات مانند وام، بیمه و مدیریت سرمایه استفاده می کند.

ZestFinance: این شرکت از تحلیل داده و هوش مصنوعی برای ارائه راهکارهای اعتباری به مشتریان استفاده می‌کند.

Credit Karma: به عنوان یک پلتفرم ارزیابی اعتبار، از تحلیل داده برای ارائه راهکارهای مالی و اعتباری استفاده می‌کند.

این شرکت‌ها به طور گسترده از تحلیل داده برای بهبود خدمات و ارائه راهکارهای مالی مبتنی بر داده استفاده می‌کنند

نتیجه‌گیری:

داده‌کاوی تأثیراتی عمیق بر شیوه‌های کسب و کار و مدیریت دانش در سال‌های اخیر داشته است. هوشمندی کسب و کار مشهورترین کاربرد فنون داده‌کاوی است. قابلیت‌های طبقه‌بندی و پیش‌بینی ابزارهای داده‌کاوی موجب کاربرد این ابزارها به منظور مقاصد پیش‌بینی و رشکستگی، وضعیت تداوم فعالیت، پیش‌بینی درماندگی، تشخیص و شناسایی تقلب مدیریت، همچنین تخمین ریسک اعتباری و عملکرد واحد تجاری شده و بدین ترتیب داده‌کاوی را تبدیل به موضوعی با درجه اهمیت زیاد در امور مالی و حسابداری نموده است. این ابزار اساسی در تصمیم‌گیری‌ها، پیش‌بینی‌ها و بهبود استراتژی‌های مالی نقش بسیار حیاتی ایفا می‌کند و برای صنعت مالی یک منبع قدرتمند و ضروری است

منابع

- دستگیر محسن، شفیعیسردشت مرتضی. فناوری داده کاوی، رویکردی نوین در حوزه مالی. دانش حسابرسی. (۱۳۹۰)
- هان، ژیاوی و کمبر، میشلن و پی، ژان، ترجمه اسماعیلی مهدی، "داده کاوی مفاهیم و تکنیک ها"، انتشارات نیاز دانش، (1393)
- سیدنژاد فهیم، سیدرضا و غلامی گلسفید، فاطمه و حسین نیا دیلمان، خدیجه، تاثیر داده های بزرگ، تجزیه و تحلیل داده ها و هوش مصنوعی بر حوزه مالی، هشتمین کنفرانس بین المللی و ملی مطالعات مدیریت، حسابداری و حقوق، تهران (1402)
- نجات، امیررضا و آرش علی اکبری (۱۳۸۷). داده کاوی راهی به سوی ناشناخته ها. ماهنامه توسعه انسانی پلیس، شماره ۱۸
- احمدلو، یعقوب، "استفاده از داده کاوی بعنوان ابزار رقابتی در صنعت بانکداری" دومین کنفرانس جهانی بانکداری الکترونیکی، تهران، مردادماه (۱۳۸۷)
- خطیب سمنانی، محمد علی، ایزدی، سمیه، خادمی، سیدرضا، (۱۳۹۶). مزیت رقابتی داده کاوی در بانکداری الکترونیکی در برابر نااطمینانی های اقتصادی، فصلنامه اقتصاد مالی، شماره ۳۹

- Kahya, Esra, Uyar ,Kumru, Ozyirmidokuz, Musta Hakan Ozyirmidokuz, " A Data Mining Based Approach to a Firm's Marketing Channel", Economics and Finance, Volume 27, (2015), Pages 77–84
- <https://khavarzadeh.com/articles>
- <https://hamrah.academy/>
- <https://a1finance.ir/>
- <https://karokasb.org/>