



الف	معرفی طرح
۱	عنوان طرح پیشنهادی ارائه رویکردی بمنظور استخراج مغایریت داده های عملیاتی با نمونه های استاندارد در اسناد مالی بانک با بهره گیری از اتوماتای یادگیر
2	شناسایی و مدیریت ریسک ■ شناسایی رفتارهای غیرمتعارف ■ تشخیص تهدیدات و بهبود پروتکل های امنیتی ■ تحلیل کانالهای ارائه خدمات ■ محور طرح پیشنهادی
3	<p>امروزه گسترش روزافزون حجم داده های عملیاتی در نظام بانکی از جمله اسناد مربوط به تراکنش ها و مبادلات مالی مشتریان در حوزه های مختلف و در قالب خدمات قابل ارائه توسط بانک ها از یک سو و همچنین ضرورت رعایت استانداردها، قواعد و دستورالعمل های درون سازمانی (نظارتی) و برون سازمانی (حاکمیتی و بین المللی) که متضمن بقاء و بالندگی بانک هاست، لزوم بهره گیری از رویکردها و تکنولوژی های نوین تحلیل داده از جمله الگوریتم های هوش مصنوعی، بمنظور سهولت و کارآمد سازی متدهای فرآوری داده های عملیاتی بمنظور استخراج اطلاعات مورد نیاز در فرآیند های کنترلی و نظارتی بیش از پیش احساس می گردد.</p> <p>نیازمندی که در صورت حصول ضمن تضمین رعایت اصول و قواعد ناظر بر فعالیت های مالی قابل ارائه و بطبع آن جلوگیری از ضرر و زیان های مالی ناشی از عدم اجرای صحیح و موفقیت آمیز اعمال نظارت کارآمد، میتواند از تبعات و خسارت های معنوی جبران ناپذیر حاصل از این خلاء برای بانک در سطح اجتماعی بطور جدی جلوگیری نماید.</p> <p>هوش مصنوعی با مفهوم بستری بمنظور ارائه الگوریتم های متنوع تحلیل داده های حاصل از تعاملات، رفتارها (مبتنی بر اندیشه و تعقل بر پایه دانش، مبتنی بر احساس و عواطف بر پایه کنش و واکنش انسانی و ...)، مفاهیم و استانداردهای تحلیلی بمنظور شبیه سازی مجموعه فعالیت های موجودات در چهارچوبی قابل فهم و تحلیل برای ماشین تحولی شگرف در زندگی انسانها رقم زده است.</p> <p>در میان انبوه الگوریتم ها و استانداردهای پرکاربرد هوش مصنوعی در زمینه تجزیه تحلیل و پردازش داده های عملیاتی با هدف تطبیق با الگوهای استاندارد، میتوان به الگوریتم های یادگیری ماشین (Machine Learning) اشاره نمود.</p> <p>الگوریتم های یادگیری ماشین یا ماشین لرنینگ این امکان را برای ماشین فراهم مینمایند تا بصورت اتوماتیک و براساس مجموعه ای شامل داده یا مشاهدات معین یادگیر بگیرد و منطق خود را بسازد و درجهت پیشرفت این یادگیری را به روز رسانی نماید.</p> <p>یادگیری تقویتی یکی از انواع یادگیری ماشین است در این نوع یادگیری با عاملی روبرو هستیم که از طریق سعی و خطا با محیط تعامل کرده و یاد میگیرد تا عملی بهینه را برای رسیدن به هدف انتخاب نماید.</p> <p>یکی از روشهای یادگیری تقویتی، اتوماتای یادگیر است. اتوماتای یادگیر با در نظر گرفتن احتمال یکسان برای تمامی اقدامهای خود برای یافتن پاسخ مسئله شروع میشود. یک اتوماتای یادگیر را میتوان به صورت یک شیء مجرد (Object Abstract) که دارای تعداد متناهی عمل، نمونه یا مشاهده می باشد در نظر گرفت. اتوماتای یادگیر با انتخاب و اعمال یک واحد عمل به محیط (P) پاسخ خود را دریافت می نماید و متناسب با پاسخ محیط به واحد عمل انتخاب شده جریمه یا پاداش اختصاص میدهد و سپس واحد عمل بعدی را بر اساس نتایج قبلی</p>



کاربرگ معرفی طرح

رویداد دیتاتون بانک سپه

<p>برای ارزیابی انتخاب و اعمال می نماید. در طی این فرایند آتوماتا یاد میگیرد که عمل بهینه را انتخاب نماید. پرواضح است میزان موفقیت و کارآمدی این الگوریتم های در تحلیل، فرآوری، تصمیم گیری در خصوص انسجام بخشی و تطبیق نمونه های استاندارد وابسته به روش های و تکنولوژی های مورد استفاده در خوانش صحیح و نمونه برداری دقیق از داده های عملیاتی بصورت هوشمند Optical Character Recognition (OCR) بمنظور تطابق با استاندارد ها، و نیز برخورداری از مجموعه ای متنوع از نمونه های استاندارد (محیط پویا) بمنظور مقایسه با نمونه های عملیاتی توسط اتوماتای یادگیر می باشد.</p> <p>هر اندازه میزان تطابق از حیث نوع و محتوای داده های خوانده شده به داده های محیط عملیاتی بیشتر باشد، و مجموعه نمونه های استاندارد متنوع تر باشد میتوان نتایج مطلوبتری از اعمال الگوریتم ها انتظار داشت.</p> <p>ما در رویکرد پیشنهادی خود سعی داریم ابتدا یک خوانش صحیح و دقیقی از محتوای اسناد مالی با بهره گیری از تکنولوژی OCR ایجاد و دسته بندی نمائیم و سپس خروجی این مرحله را بصورت نمونه های عملیاتی دسته بندی شده بر اساس محتوا و نوع و کاربرد، را با استفاده از الگوریتم آتوماتای یادگیر با نمونه های استاندارد متناظر (محیط) و هم نوع مقایسه می کنیم و متناسب با میزان شباهت با نمونه استاندارد نسبت به جریمه یا پاداش دهی به نمونه عملیاتی مورد نظر اقدام می نمائیم.</p> <p>در پایان هر دور مقایسه نسبت به دسته بندی مجدد نمونه های اقدام می گرد و سپس مقایسه با دسته های همجوار بر اساس میزان همسایگی صورت پذیرفته و امتیاز دهی مجدد انجام میگردد، در پایان مقایسه نمونه های عملیاتی دارای بیشترین جریمه در هر مرحله استخراج و در صورت عدم کسب امتیاز پاداش مناسب در هیچ یک از مراحل بعنوان نمونه مغایر برجسب گذاری میشوند، نمونه های که بالاترین امتیاز پاداش در هر مرحله را دارند بعنوان نمونه استاندارد گروه مطبوع در کنار نمونه های موجود در نظر گرفته میشوند تا دامنه نمونه های عملیاتی استاندارد (محیط) به روزرسانی و افزایش یابد. انتظار می رود در صورت انجام دقیق هرکدام از مراحل نتایج مطلوبی در تشخیص و استخراج مغایرت میان نمونه های عملیاتی اسناد و قرار دادها با استاندارد ها و دستورالعمل ها حاصل گردد.</p>		
<p>انتظار می رود در صورت انجام دقیق هرکدام از مراحل نتایج مطلوبی در تشخیص و استخراج مغایرت میان نمونه های عملیاتی اسناد و قرار دادها با استاندارد ها و دستورالعمل ها حاصل گردد. از این رو سودآوری مطلوبی از رهگذر صرفه جویی در زمان و هزینه انجام کنترل های سنتی و غیر هوشمند برای نهادهای مالی حاصل میگردد. همچنین با توجه به میزان موفقیت مورد انتظار در فرآیندی های کنترلی حاصل از پیاده سازی رویکرد مورد نظر به میزان قابل ملاحظه ای قابلیت اطمینان و اعتماد مشتریان و نهادهای نظارتی به سلامت فرآیندهای کنترلی در نهاد مالی افزایش خواهد یافت.</p> <p>پر واضح است که با توجه به میزان موفقیت مورد انتظار از عملکرد صحیح رویکرد پیشنهادی در استخراج مغایرت های احتمالی، میزان ضرر های مادی و معنوی ناشی از عدم تشخیص مغایرت و سوء استفاده های احتمالی به میزان زیادی کاهش خواهد یافت.</p>	<p>منافع و جنبه های کاربردی پذیری</p>	<p>4</p>
<p>Optical Character Recognition (OCR) Learning Automaton (LA)</p>	<p>تعیین الگوریتم های مورد استفاده</p>	<p>5</p>



کاربرگ معرفی طرح

رویداد دیتاتون بانک سپه

<p>تکنولوژی (Optical Character Recognition (OCR بمنظور خوانش داده های عملیاتی مورد استفاده قرار میگیرد. درصد موفقیت در خوانش صحیح و مطابقت بالای داده های خوانده شده با واقعیت مندرج در اسناد فیزیکی ارتباط مستقیمی با میزان موفقیت در تصمیم گیری ، انسجام بخشی و مقایسه با نمونه های استاندارد دارد.</p> <p>آتوماتای یادگیر (LA) ماشینی است که با انتخاب یک عمل و ارزیابی آن با محیط پیرامون (مجموعه ای متناهی از نمونه ها) نسبت به جریمه یا پاداش عمل اقدام می نماید بدین صورت که اگر عمل مورد نظر بر اساس معیارهای تعیین شده محیط پیرامون مطابقت داشت پاداش و در غیر این صورت جریمه میشود و سپس ماشین نمونه یا عمل بعدی را جهت مطابقت با محیط انتخاب می نماید.</p>		
<p>Python ■ Matlab ■ .Net ■</p>	<p>نرم افزارهای مورد نیاز</p>	<p>6</p>
<p>داده های مورد نیاز بمنظور پیاده سازی، ارزیابی و صحت سنجی و اجرایی نمودن این رویکرد شامل داده های عملیاتی اسناد و قراردادهای مالی در نهاد مالی بانک می باشد.</p> <p>در این رویکرد ما گام نخست ارکان اسناد مالی همچون ضمانت نامه، اعتبار اسنادی، چک و... را بمنظور صحت سنجی اصالت از حیث صحت محتوا (حقوقی و نظارتی) نظیر درج صحیح و دقیق مشخصات هویتی طرفین تبادل، امضاء و... و نیز مقداری (حقوقی و نظارتی) مانند تاریخ، مبلغ، تعداد و ... در مقایسه با نمونه های استاندارد مورد بررسی قرار میدهیم.</p>	<p>داده های مورد نیاز برای اجرای تحلیل</p>	<p>7</p>
<p>افزایش سودآوری ■ افزایش سهم بازار ■ افزایش کیفیت یا ظرفیت ■ مدیریت و کنترل ■</p>		
<p>این رویکرد از رهگذر خوانش صحیح، انسجام بخشی و مقایسه داده های عملیاتی با نمونه های استاندارد منطبق با دستورالعمل های درون و برون سازمانی ضمن حصول تسهیل و کارآمد سازی و افزایش کیفیت و ظرفیت های کنترلی متناسب با حجم داده ها، امکان اعمال مدیریت کنترل فرآیند های نظارتی در پایش ، واپایش و نظارت بر داده های عملیاتی و در نتیجه افزایش دقت، صحت و قابلیت اطمینان کنترل اسناد و قراردادهای، کاهش سربارهای هزینه ای متاثر از اجرای رویکردهای کنترلی فعلی از طریق صرفه جویی در منابع و زمان مورد نیاز در رویکردهای سنتی و غیر هوشمند، ضرر و زیان های مادی و معنوی ناشی از عدم امکان نظارت کارآمد و متناسب با حجم عظیم داده های عملیاتی را به میزان قابل توجهی بهبود می بخشد که در نهایت سودآوری حاصل از کاهش هزینه های جانبی و افزایش برخورداری و پذیرش سهم بیشتر از بازار تقاضا متاثر از افزایش قابلیت اطمینان را برای بانک محقق خواهد نمود.</p>	<p>زمینه کاربرد</p>	<p>8</p>



<p>همانطور که پیش تر بیان شد افزایش کارآمدی و انعطاف پذیری الگوریتم های پردازشی و تحلیلی مورد استفاده ما وابسته به کیفیت و صحت خوانش اسناد فیزیکی می باشد از اینرو ما از میان انبوه مقالات تحقیقاتی در این حوزه به بررسی چند نمونه پرداختیم تا با بهره گیری نتایج حاصله رویکرد بهینه در خوانش اسناد و قرار داده ها مورد استفاده قرار گیرد. همچنین در خصوص پیاده سازی الگوریتم اتوماتای یادگیر به بررسی چند نمونه از تحقیقات انجام شده در این حوزه خواهیم پرداخت.</p> <p>نوحی و همکارانش [1]، مسئله خوانش و ایجاد خودکار نمایه برای تصاویر متنی به زبان فارسی را بر اساس بهره گیری از ضریب همبستگی و نیز رویکرد گسترش تصویر بمنظور استخراج نمونه های با حداکثر شباهت ممکن مورد بررسی قرار داده اند نتایج حاصل از روش پیشنهادی آن ها موفقیت 90 درصدی به همراه دامنه خطایی در حدود 6.25 برای False Negative (FN) و 3.75 برای False Positive (FP)، در خوانش و نمایه گذاری را ارائه نموده است.</p> <p>فرامرزی [2]، مسئله خوانش و تشخیص خودکار صفحات فهرست با توجه به الگوی آنها در پایان نامه های فارسی و لاتین را بر روی پایگاه اطلاعاتی پژوهشگاه اطلاعات و مدارک علمی ایران شامل پایان نامه های کارشناسی ارشد و دکترای دفاع شده در دانشگاه های داخل کشور از سال 1310 مورد ارزیابی قرار داده است، رویکردی پیشنهادی در این تحقیق مبتنی بر رویکرد پیشنهادی (Mandal, et, al, 2003) می باشد، برای اینکه این رویکرد نتایج قابل قبولی برای متون فارسی ایجاد نماید نسبت به پایش متون اقدام نموده و متون را از نظر پاک سازی (حذف زواید تزئینی) و فریم بندی بازنگری نموده اند، اعمال رویکرد مورد نظر بر روی متن پایش شده موفقیتی در حدود 99.7 درصد ارائه نموده است. دامنه خطایی این رویکرد در حدود 0.06 درصد برای False Positive (FP)، بیان شده است.</p> <p>موگویی و همکارانش [3]، مسئله بهینه سازی هزینه تخصیص و بازپس گیری منابع در محیط رایانش ابری را مورد بررسی قرار داده اند، آنها توانسته با بهره گیری از اتوماتای یادگیر هزینه تخصیص و بازپس گیری منابع را با محوریت میزان شباهت منبع (سرعت، فضا و...) به درخواست تخصیص، بهبود دهند.</p> <p>حسینی سدهی و همکارانش [4]، مسئله تشخیص لبه تصاویر را با بهره گیری از اتوماتای یادگیر سلولی دو بعدی با R سطر و C ستون که در هر سلول یک اتوماتای یادگیر با ساختار ثابت وجود دارد را مورد بررسی قرار داده اند. آنها رویکرد خود را با دو روش تشخیص لبه کنی و سو بل مقایسه نموده اند، روش پیشنهادی آنها از دقت بیشتری در تشخیص برخوردار است چراکه در هر همسایگی از شدت روشنایی سلولهای موجود در همسایگی برای تعیین لبه استفاده شده است.</p>	<p>9 نمونه کارهای مشابه</p>
---	-----------------------------



کاربرگ معرفی طرح

رویداد دیتاتون بانک سپه

10	سایر موارد	<p>منابع و مآخذ :</p> <p>[1] نوحی علیرضا، یغمایی فرزین، ایجاد خودکار نمایه برای تصاویر متنی به زبان فارسی ، نخستین کنفرانس بین المللی پردازش خط و زبان، 1391.</p> <p>[2] فرامرزی اسماعیل، و تشخیص خودکار صفحات فهرست با توجه به الگوی آنها در پایان نامه های فارسی و لاتین، نشریه پردازش و مدیریت اطلاعات مدیریت، 1385، 46، 34-1.</p> <p>[3] موگویی خسرو، قبايي آرانی مصطفی، شمسی محبوبه، رویکردی مبتنی بر هزینه برای مقیاس بندی اتوماتیک منابع در محیط رایانش ابری ، هفتمین کنفرانس ملی مهندسی برق و الکترونیک ایران، 1394.</p> <p>[4] حسینی سدهی سید میثم و مبینی محمدرضا، تشخیص لبه در تصاویر با استفاده از اتوماتای یادگیر سلولی، سیزدهمین کنفرانس انجمن ملی کامپیوتر ایران، 1386.</p>
----	------------	--

فرم (2)

ب	اعضای تیم						
نام	نام خانوادگی	سمت سازمانی	محل خدمت	آخرین مدرک تحصیلی	نقش در تیم	شماره تلفن همراه	پست الکترونیکی
1	وحید	سلیمانی	معاون	بازرسی غیر حضوری	کارشناسی ارشد	سرپرست	09128939107
2	خسرو	موگویی	بازرس	بازرسی غیر حضوری	کارشناسی ارشد	کارشناس داده تحلیل و پیاده سازی	09386025533 KHOSRO_MOGOUIE@YAHOO.COM
3	امیر	باقری	بازرس	بازرسی غیر حضوری	کارشناسی	تحلیل و پیاده سازی	09194859227
4	نجمه	خوش مشربان	بازرس	بازرسی غیر حضوری	کارشناسی ارشد	کارشناس داده	09124786846



کاربرگ معرفی طرح

رویداد دیتاتون بانک سپه

سایر ملاحظات:

- لطفاً در صورتی که پاسخ به برخی از بندها مقدور نمی‌باشد آن را خالی بگذارید؛ هر چند تکمیل بودن کاربرگ در نظر هیئت داوران تأثیرگذار خواهد بود.
- فایل خام این کاربرگ بر روی سایت شرکت نوآوران بوم گستر امید (امیدبوم) به نشانی <https://omidinofactory.ir/datathon> قرار دارد. شرکت کنندگان بایستی پس از بارگیری فایل فرم و تکمیل آن، فایل را در همان سامانه بارگذاری نمایند.
- مهلت ارسال کاربرگ‌های تکمیل شده **30 مهرماه 1403** می‌باشد.
- در صورت وجود مشکل با شماره پشتیبانی سایت شماره 09120913915 تماس حاصل فرمایید و جهت اخذ اطلاعات تکمیلی به کانال روبیکایی با نام

مراجعه فرمایند.



"دیتاتون بانک سپه" با کد QR