

## Designing a smart model for granting banking facilities based on big data based on demographic characteristics, business status and past payments of customers

Masoumeh Vakili	Department of Information Technology Management, SR.C., Islamic Azad University, Tehran, Iran.
Maryam Rahmaty*	Department of Management, Cha.C., Islamic Azad University, Chalus, Iran.
Seyed Ahmad Shayannia	Department of Industrial Management, Fi.C., Islamic Azad University, Firoozkooh, Iran.
Reza Radfar	Department of Industrial Management, SR.C., Islamic Azad University, Tehran, Iran.

### Abstract

Designing a smart model for granting banking facilities can reduce many of the concerns and worries of the banking system regarding the repayment of granted loans and the increase in non-performing loans. For this purpose, it is necessary to design such a model using artificial intelligence algorithms. In this research, the aim is to design a smart model for granting banking facilities based on big data and based on demographic characteristics, business status and past payments of customers. The model consisted of 12 variables, the data of which was collected from the banking system and then implemented with the help of machine learning algorithms. Three algorithms, nearest neighbor, decision tree and random forest, were used in the present study, and the results show that the decision tree algorithm is able to predict customer repayment behavior with an accuracy of 99.6% at the highest level, and this indicates that the 12 variables introduced can classify customer repayment based on three categories of customers with high, medium and low risk with more than 99%.

**Keywords:** Intelligent model - Bank facility granting - Big data - Demographic characteristics - Business status - Past payments - Machine learning

**How to Cite:** Vakili, M. , Rahmaty, M. , Shayannia, S. A. and Radfar, R. (2025). Designing a smart model for granting banking facilities based on big data based on demographic characteristics, business status and past payments of customers. Journal of Intelligent Strategic Management .4(3), 613-632.

doi: bumara.3.2.1123.365.28863



Intelligent Strategic Management (JISM) in Development and Evolution is licensed under a Creative Commons Attribution-Non Commercial 4.0 International License.

© Authors

\* Corresponding Author : Rahmaty.Maryam@iau.ac.ir

## طراحی مدل هوشمند اعطای تسهیلات بانکی مبتنی بر کلان داده بر اساس مشخصه های جمعیت شناختی، وضعیت کسب و کار و پرداختهای گذشته مشتریان

گروه مدیریت فناوری اطلاعات، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران.

معصومه وکیلی

گروه مدیریت، واحد چالوس، دانشگاه آزاد اسلامی، چالوس، ایران.

مریم رحمتی\*

گروه مدیریت صنعتی، واحد فیروز کوه، دانشگاه آزاد اسلامی، فیروز کوه، ایران.

سید احمد شایان نیا

گروه مدیریت صنعتی، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران.

رضا رادفر

### چکیده

طراحی یک مدل هوشمند اعطای تسهیلات بانکی می تواند بسیار از دغدغه ها و نگرانیهای سیستم بانکی را در خصوص بازپرداخت وامهای اعطای شده و افزایش وامهای غیر عملیاتی را کاهش دهد. برای این منظور لازم است با بهره گیری از الگوریتمهای هوش مصنوعی چنین مدلی طراحی شود. در این تحقیق هدف طراحی یک مدل اعطای تسهیلات بانکی هوشمند بر اساس کلان داده و بر اساس مشخصه های جمعیت شناختی وضعیت کسب و کار و پرداختهای گذشته مشتریان می باشد. مدل شامل ۱۲ متغیر بود که داده های آن از سیستم بانکی جمع اوری شده و سپس با کمک الگوریتمهای یادگیری ماشین پیاده سازی شد. سه الگوریتم کاندیدترین همسایه، درخت تصمیم و جنگل تصادفی در تحقیق حاضر استفاده شد که نتایج نشان می دهد الگوریتم درخت تصمیم با دقت ۹۹,۶ درصد در بالاترین سطح قادر به پیش بینی رفتار بازپرداخت مشتریان بوده و این امر نشان می دهد که متغیرهای ۱۲ گانه معرفی شده تا بیش از ۹۹ درصد می توانند بازپرداخت مشتریان را بر اساس سه دسته مشتریان با ریسک بالا، متوسط و ریسک اندک دسته بندی نمایند.

**کلیدواژه‌ها:** مدل هوشمند- اعطایی تسهیلات بانکی- کلان داده- مشخصه های جمعیت شناختی- وضعیت کسب و کار- پرداختهای گذشته- یادگیری ماشین

**استناد به این مقاله:** وکیلی، معصومه و رحمتی، مریم و شایان نیا، سید احمد و رادفر، رضا. (۱۴۰۴). طراحی مدل هوشمند اعطای تسهیلات بانکی مبتنی بر کلان داده بر اساس مشخصه های جمعیت شناختی، وضعیت کسب و کار و پرداختهای گذشته مشتریان. مدیریت استراتژیک هوشمند، ۴(۳)، ۶۳۲-۶۱۳.



مدیریت استراتژیک هوشمند (JISM) در توسعه و تکامل تحت مجوز بین المللی کرییتیو کامنز با شرایط انتساب-غیرتجاری ۴,۰ منتشر می شود.

©نویسندگان

\* نویسنده مسئول: Rahmaty.Maryam@iau.ac.ir

## مقدمه

یکی از دغدغه های مهم سیستم های بانکی و مدیران آنها شناسائی متقاضیان وامی می باشد که قادر به بازپرداخت وام بانکها باشند چرا که بانکها از یک سو نیازمند جذب سپرده و از سوی دیگر نیازمند اعطای وام هستند تا بتوانند به سودآوری دست یافته و جریان نقدینگی موجود را به سمت و سوی مناسب سوق دهند (هوتا و همکاران، ۲۰۲۵). اما عدم بازپرداخت یا قصور در بازپرداخت می تواند به افزایش وامهای غیر عملیاتی بانکها و در نهایت ناترازی و ورشکستگی بانکها منجر شود و این موضوعی است که کشور به وفور مشاهده می شود (بی و بائو، ۲۰۲۴). بانکها نیازمند اعطای وام هستند اما همواره به دنبال این هستند که افرادی را شناسایی کنند که در بازپرداخت وام مشکلی ایجاد نمی کنند و این در واقع در زمره اولویتهای اصلی آنان می باشد. افزایش وامهای غیر عملیاتی به سبب اعطای وام به افرادی که در بازپرداخت ضعیف هستند می تواند منجر به ناکارائی بانکها شده و مدیران بانکها همواره از این موضوع هراس دارند (زوآما و همکاران، ۲۰۲۴).

اما تشخیص اینکه اعطای تسهیلات باید به چه کسانی صورت گیرد امر ساده ای نیست و همواره یک مشکل و مسئله تلقی می شود (الرشیدی و همکاران، ۲۰۲۴). در این خصوص تحقیقات بسیاری انجام شده که در آن تلاش شده مدلی برای اعطای تسهیلات بانکی طراحی شود. در برخی از این تحقیقات از الگوریتمهای یادگیری ماشین استفاده شده است این الگوریتم ها به سبب ماهیت خاص خود قادر به تشخیص و دسته بندی پدیده ها می باشند و لذا با وجود حجم بالای کلان داده که یکی از ویژگیهای مهم سیستم بانکی می باشد امکان پردازش این کلان داده با کمک الگوریتمهای یادگیری ماشین وجود دارد (سیناپ، ۲۰۲۴). آنان با ورود داده ها و آموزش آنها قادر به دسته بندی مشتریان از نظر سطح ریسک بازپرداخت بر اساس داده های گذشته می باشند (عبدالهادی و همکاران، ۲۰۲۵).

اما مسئله اصلی این است که داده های گذشته صرفه نمی تواند پیش بینی کننده بازپرداخت وام باشد و بسیاری از سیستمهای هوشمند تلاش کرده اند بر اساس داده های گذشته به پیش بینی بازپرداخت پردازند که نتایج چندان دقیقی از این موضوع حاصل نشده است. لذا متغیرها و عوامل دیگری نیز باید بررسی شوند که مهمترین آنها می تواند وضعیت کسب و کار متقاضی و همچنین ویژگیهای جمعیت شناختی باشد. البته ویژگیهای جمعیت

شناختی نیز در برخی تحقیقات مورد توجه قرار گرفته اما ویژگیهای کسب و کار در اغلب تحقیقات مغفول مانده است که تحقیق حاضر به دنبال رفع آن می باشد. در واقع تحقیق حاضر تلاش دارد مدل جدیدی را در خصوص سیستم اعطای تسهیلات بانکی طراحی نماید که بر اساس کلان داده عمل کرده و مبتنی بر سه دسته متغیر اصلی یعنی متغیرهای جمعیت شناختی، متغیرهای مربوط به کسب و کار و متغیرهای مربوط به سابقه پرداخت می باشد این متغیرها که مشتمل بر ۱۲ مورد می باشد در تحقیقات پیشین در نظر گرفته نشده و تحقیق حاضر از این نظر مشمول نوآوری می باشد. مدل بر اساس سه الگوریتم درخت تصمیم، جنگل تصادفی و کا نزدیکترین همسایه پیاده سازی شده و بر این اساس مشتریان به سه دسته کم ریسک، پر ریسک و با ریسک متوسط در خصوص بازپرداخت وام تفکیک می شوند که بر این اساس مبنای تصمیم گری مدیران در خصوص اعطای وام خواهد بود.

ساختار مقاله حاضر به این صورت است که در بخش بعدی مرور ادبیات ارائه شده و سپس شکاف تحقیقاتی بر اساس آن استخراج می شود در ادامه روش شناسی ارائه گردیده و الگوریتم ها معرفی می گردند. پس از آن تجزیه و تحلیل یافته ها صورت گرفته و در انتها نتایج تبیین و تشریح می گردند.

### مرور ادبیات

در این بخش به مروری بر تحقیقات انجام شده در حوزه طراح مدل هوشمند اعطای تسهیلات وام پرداخته می شود. تمرکز اغلب تحقیقات در حوزه یادگیری ماشین و بهره گیری از آن در جهت طراحی مدل اعطای تسهیلات به شکل هوشمند می باشد. دانسانا و همکاران (۲۰۲۳) به تحلیل اثر ویژگیهای وام بر پیش بینی وام بانکی با استفاده از الگوریتم جنگل تصادفی می پردازند. راقاوان (۲۰۲۴) به مدیریت داده بنیان برای پردازش وام هوشمند می پردازد. زوآما<sup>۲</sup> و همکاران (۲۰۲۴) به پیاده سازی الگوریتمهای یادگیری ماشین برای پیش بینی قصور در بازپرداخت وام بر اساس رفتار مشتری توجه دارد. حق<sup>۳</sup> و همکاران (۲۰۲۴) پیش بینی وام بانکی را با استفاده از تکنیکهای یادگیری ماشین انجام می دهند. حسین و همکاران (۲۰۲۴) سیستمی را جهت پیش بینی وام بانکی با استفاده از مدلهای یادگیری ماشین ارائه می کنند. ال گیمی<sup>۴</sup> و همکاران (۲۰۲۴) به ارتقای خدمات

<sup>1</sup> Dansana et al

<sup>2</sup> Prathap Raghavan

<sup>3</sup> Zuama et al

<sup>4</sup> Haque

<sup>5</sup> EL-GEMEIE et al

بانکی از طریق چارچوب تحلیل داده های هوشمند توجه داشته و مدلی در این خصوص ارائه می کنند.

هریتسنکو و همکاران (۲۰۲۴) مروری بر تکنولوژیهای هوشمند در بانکداری در یک مقاله مروری دارند. حسن (۲۰۲۴) یک سیستم پیش بینی برانندگی وام را بر اساس تکنیکهای یادگیری ماشین طراحی می کنند که از تکنیک انجمنی با تقاضای هوشمند برای طرحهای زیرساختی بهره می گیرد. الرشیدی<sup>۱</sup> و همکاران (۲۰۲۴) سیستم پشتیبانی از تصمیم هوشمند را برای ارزیابی وام با استفاده از یادگیری ماشین طراحی و پیاده سازی می کنند. سیناپ (۲۰۲۴) به مطالعه تطبیقی پیش بینی تائید وام با استفاده از روشهای یادگیری ماشین اقدام می کنند. فاتی<sup>۲</sup> و همکاران (۲۰۲۴) یک مدل پیش بینی ناتوانی در بازپرداخت وام را با استفاده از یادگیری ماشین و مهندسی مشخصه ها ارائه می کنند. چادوری<sup>۳</sup> و همکاران (۲۰۲۴) از هوش مصنوعی برای ساده سازی پرداختها و وامدهی برای تحقق یک اکو سیستم مالی ساده تر بهره می گیرند. بی و بائو<sup>۴</sup> (۲۰۲۴) از تکنولوژی هوش مصنوعی در مدیریت ریسک اعتبار بانکی استفاده می کند. حسین<sup>۵</sup> (۲۰۲۵) به تحلیل تطبیقی مدل‌های یادگیری ماشین برای پیش بینی ریسک اعتبار در سیستمهای بانکی می پردازد. سریرام<sup>۶</sup> و همکاران (۲۰۲۵) از یادگیری ماشین و هوش مصنوعی برای مدل‌های ارزیابی ریسک اعتبار نسل بعدی بهره می گیرند. عبدالهادی<sup>۷</sup> و همکاران (۲۰۲۵) یک مدل داده محور را برای پیش بینی تائید وام با استفاده از رویکردهای یادگیری ماشین ارائه می کنند. تان و همکاران (۲۰۲۵) پیش بینی ریسک اعتبار را با شاخص درک فساد با استفاده از رویکردهای یادگیری ماشین انجام می دهند. هوتا و همکاران (۲۰۲۵) به ارزیابی عملکرد پیش بینی تائید وام در بخش مالی می پردازند.

---

<sup>1</sup> Hassan

<sup>2</sup> Elrashidy et al

<sup>3</sup> Vahid SINAPI

<sup>4</sup> Fati et al

<sup>5</sup> Chaudhry

<sup>6</sup> Shuo Chen Bi 1, a, Wenqing Bao

<sup>7</sup> Hossain

<sup>8</sup> Sriram

<sup>9</sup> Abdelhady

## جدول ۱: مرور ادبیات

محققین	سال	هدف	مدل هوشمند اعطای تسهیلات بانکی	مشخصه های جمعیت شناختی	وضعیت کسب و کار	پروژه های گذشته	یادگیری ماشین
دانسانا و همکاران	2023	تحلیل اثر ویژگیهای وام بر پیش بینی وام بانکی با استفاده از الگوریتم جنگل تصادفی	<input type="checkbox"/>	<input type="checkbox"/>			<input type="checkbox"/>
راقاوان	2024	مدیریت داده بنیاد پردازش وام	<input type="checkbox"/>				
زواما و همکاران	2024	پیاده سازی الگوریتمهای یادگیری ماشین برای پیش بینی قصور در بازپرداخت وام بر اساس رفتار مشتری	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>
حق و همکاران	2024	پیش بینی وام بانکی با استفاده از تکنیکهای یادگیری ماشین	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>
حسین و همکاران	2024	پیش بینی وام بانکی با استفاده از مدل های یادگیری ماشین	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>
ال گیمی و همکاران	2024	ارتقای خدمات بانکی از طریق چارچوب تحلیل داده های هوشمند	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>
هریتسنکو و همکاران	2024	مروری بر تکنولوژیهای هوشمند در بانکداری	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>
حسن	2024	طراحی سیستم پیش بینی برآوردگی وام بر اساس تکنیکهای یادگیری ماشین	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>
الرشیدی و همکاران	2024	طراحی سیستم پشتیبانی از تصمیم هوشمند برای ارزیابی وام با استفاده از یادگیری ماشین	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>
سیناپ	2024	مطالعه تطبیقی پیش بینی تأیید وام با استفاده از روشهای یادگیری ماشین	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>

محققین	سال	هدف	مدل هوشمند اعطای تسهیلات بانکی	مشخصه های جمعیت شناختی	وضعیت کسب و کار	پرداختهای گذشته	یادگیری ماشین
فاتی و همکاران	2024	طراحی مدل پیش بینی ناتوانی در بازپرداخت وام با استفاده از یادگیری ماشین و مهندسی مشخصه ها	<input type="checkbox"/>				<input type="checkbox"/>
چادوری و همکاران	2024	استفاده از هوش مصنوعی برای ساده سازی پرداختها و وامدهی به منظور تحقق یک اکو سیستم مالی ساده تر	<input type="checkbox"/>				<input type="checkbox"/>
بی و باثو	2024	استفاده از تکنولوژی هوش مصنوعی در مدیریت ریسک اعتبار بانکی	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>
حسین	2025	تحلیل تطبیقی مدل‌های یادگیری ماشین برای پیش بینی ریسک اعتبار در سیستمهای بانکی	<input type="checkbox"/>				<input type="checkbox"/>
عبدالهادی و همکاران	2025	ارائه مدل داده محور برای پیش بینی تائید وام	<input type="checkbox"/>				<input type="checkbox"/>
تان و همکاران	2025	پیش بینی ریسک اعتبار با شاخص درک فساد با استفاده از رویکردهای یادگیری ماشین	<input type="checkbox"/>			<input type="checkbox"/>	<input type="checkbox"/>
هوتا و همکاران	2025	ارزیابی عملکرد پیش بینی تائید وام در بخش مالی	<input type="checkbox"/>				<input type="checkbox"/>
تحقیق حاضر		طراحی مدل هوشمند اعطای تسهیلات بانکی مبتنی بر کلان داده بر اساس مشخصه های جمعیت شناختی، وضعیت کسب و کار و پرداختهای گذشته مشتریان	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

بر اساس جدول فوق می توان مشاهده کرد که تحقیقات بسیاری در حوزه ارائه مدل هوشمند اعطای تسهیلات بانک صورت گرفته است. برخی از این مدلها به پیش بینی ناتوانی در بازپرداخت اقدام کرده و برخی نیز پیش بینی بازپرداخت وام را انجام داده اند در این تحقیقات تمرکز بر یادگیری ماشین و بهره گیری از این الگوریتم ها می باشد که عمده مدلهای ارائه شده مبتنی بر سابقه پرداختهای گذشته و رفتار مشتری می باشد این در حالیست که عوامل دیگری می تواند تعیین کننده در بازپرداخت وام یا ناتوانی در آن ایفای نقش نماید که در تحقیقات پیشین مورد توجه قرار نگرفته است عمده این متغیرها شامل مشخصه های جمعیت شناختی و وضعیت کسب و کار می باشد. که توضیح جامعتر آنها در بخش روش شناسی ارائه می گردد. در مجموع نوآوری تحقیق حاضر در ارائه یک مدل جدید برای تسهیلات هوشمند می باشد که در تحقیقات پیشین کمتر مورد توجه قرار گرفته و لذا از این نظر شکاف تحقیقاتی مشاهده می شود که تحقیق حاضر به دنبال رفع آن می باشد.

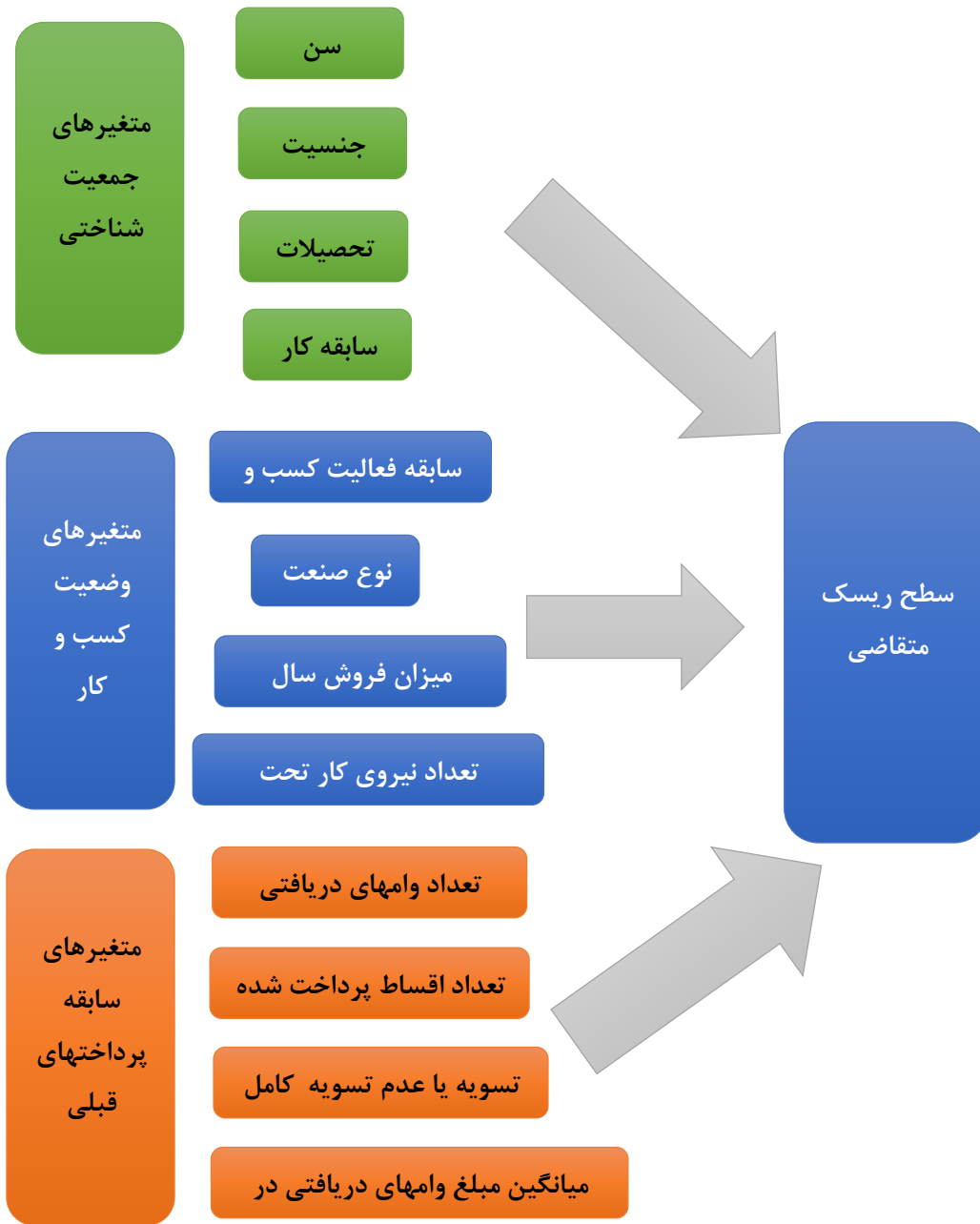
### روش شناسی

تحقیق حاضر از نظر هدف کاربردی و از نوع توسعه ای می باشد چرا که در آن از الگوریتمهای یادگیری ماشین برای پیاده سازی مدل هوشمند اعطای تسهیلات استفاده می شود. در این تحقیق متغیرهای ورودی وارد مدل شده و بر اساس آن پیش بینی وضعیت بازپرداخت کنند صورت گرفته و لذا تعیین می شود که آیا اعطای تسهیلات به متقاضی امری بهینه است یا خیر. مدل ارائه شده قادر به دسته بندی مشتریان به کم ریسک، دارای ریسک متوسط و ریسک بالا می باشد که این خروجی بر اساس متغیرهای جمعیت شناختی، وضعیت کسب و کار و پرداختهای گذشته وام تعیین می شود. البته این عوامل به عنوان ابعاد مدل پیشنهادی بوده و می توان آنها را به موارد جزئی تر تفکیک نمود. متغیرهای مورد نظر در جدول ذیل ارائه شده اند.

جدول ۲: متغیرهای تحقیق

ردیف	متغیر	نماد متغیر	نوع متغیر	مقیاس متغیر
۱	سن	X1	ورودی	کمی
۲	جنسیت	X2	ورودی	باینری
۳	تحصیلات	X3	ورودی	رده ای
۴	سابقه کار کلی	X4	ورودی	کمی
۵	سابقه فعالیت کسب و کار	X5	ورودی	کمی
۶	نوع صنعت	X6	ورودی	رده ای
۷	میزان فروش سال گذشته	X7	ورودی	کمی
۸	تعداد نیروی کار تحت استخدام	X8	ورودی	کمی
۹	تعداد وامهای دریافتی	X9	ورودی	کمی
۱۰	تعداد اقساط پرداخت نشده	X10	ورودی	کمی
۱۱	تسویه یا عدم تسویه کامل وامهای قبلی	X11	ورودی	باینری
۱۲	میانگین مبلغ وامهای دریافتی در گذشته	X12	ورودی	کمی

می توان متغیرهای جدول فوق را در قالب مدل مفهومی ذیل ترسیم نمود.



شکل ۱: مدل مفهومی تحقیق

در مدل مفهومی فوق همانگونه که مشاهده می شود سه متغیر اصلی وجود دارد که هر یک دارای ۴ زیر متغیر می باشد بر این اساس مدل ارائه شده شامل ۱۲ متغیر ورودی می باشد که می تواند مشتریان را به سه دسته پر ریسک، کم ریسک و با ریسک متوسط تقسیم

نماید. داده های مربوط به این متغیرها از دیتاست بانکها جمع اوری شده و وارد مدل می شود. داده ها مربوط به حجم بالای متقاضیان تسهیلات بانکی می باشد که در حد ۱۰ هزار رکورد بوده و این داده ها مبنای تشکیل دیتاست مورد مطالعه در تحقیق حاضر می باشد. داده ها به سه دسته آموزش، اعتبارسنجی و تست تقسیم شده که ۷۰ درصد داده ها مربوط به آموزش و ۱۵ درصد مربوط به تست و اعتبارسنجی می باشد. سپس الگوریتمهای یادگیری ماشین که بیشتر در حوزه دسته بندی مورد استفاده قرار می گیرند بر اساس این داده ها پیاده سازی شده و عمل دسته بندی را انجام می دهند. الگوریتمهای مورد استفاده به شرح ذیل هستند

۱. الگوریتم درخت تصمیم

۲. الگوریتم جنگل تصادفی

۳. الگوریتم کانزدیکترین همسایه

عملکرد این الگوریتم ها بر اساس چهار معیار ذیل تعیین می شود

۱. Accuracy

۲. Precision

۳. Recall

۴. F1score

هر چه امتیاز حاصل از الگوریتم ها در خصوص این معیارها بالاتر باشد می توان گفت الگوریتم مربوطه از کارایی بالاتری برخوردار می باشد accuracy نشانگر تعداد نمونه های دسته بندی شده به صورت صحیح نسبت به کل داده های نمونه می باشد فرمول محاسبه آن به شرح ذیل است

$$accuracy = \frac{TN + TP}{TN + FP + TP + FN} \quad (1)$$

در رابطه فوق

TN کل موارد منفی صحیح

TP کل موارد مثبت صحیح

FP کل موارد منفی مثبت

FN کل موارد منفی غلط می باشد

Precision نشانگر مقدار پیشگویانه مثبت در دسته بندی نمونه داده ها می باشد فرمول

آن به شرح ذیل است

$$Precision = \frac{TP}{FP + TP} \quad (۲)$$

معیار بعدی recall می باشد که به عنوان حساسیت یا نرخ مثبت واقعی تعریف می شود فرمول آن به شرح ذیل است

$$Recall = \frac{TP}{FN + TP} \quad (۳)$$

و در نهایت معیار نهایی ارزیابی کارایی الگوریتمهای یادگیری ماشین در دسته بندی F1 Score می باشد که هم precision و هم recall را به شکل توام محاسبه کرده و به شرح ذیل می باشد.

$$F1 \text{ score} = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (۴)$$

در فرمول فوق precision در صورت در recall ضرب می شود اما در معراج این دو معیار با هم جمع شده و در ۲ ضرب می شود که نتیجه آن مقدار f1 score بوده و هر چه بالاتر باشد نشانگر کارایی بهتر یک الگوریتم است.

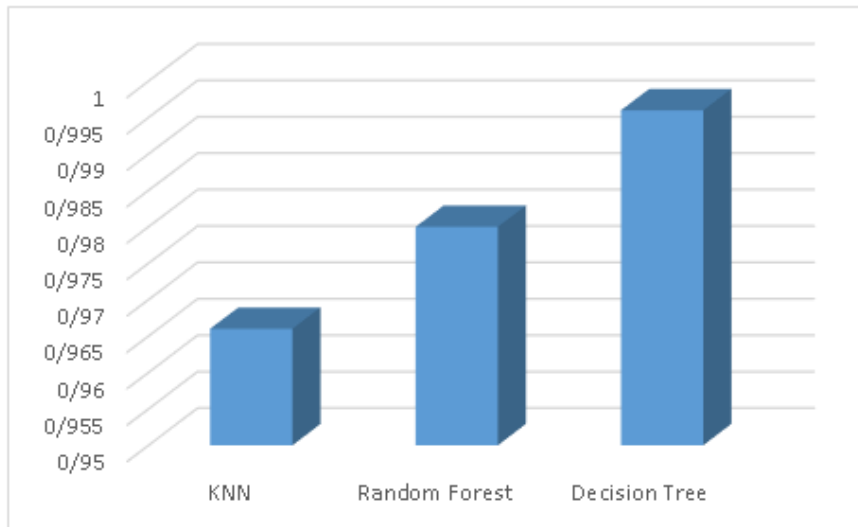
### تجزیه و تحلیل

در این بخش به ارائه نتایج حاصل از پیاده سازی الگوریتم ها و ارزیابی آنها ارائه می شود در ابتدا مقایسه الگوریتم ها از نظر چهار معیار تشریح شده در بخش روش شناسی صورت می گیرد. نتایج در جدول ۳ خلاصه شده است.

جدول ۳: مقایسه الگوریتم ها از نظر چهار معیار accuracy, precision, recall, f1 score

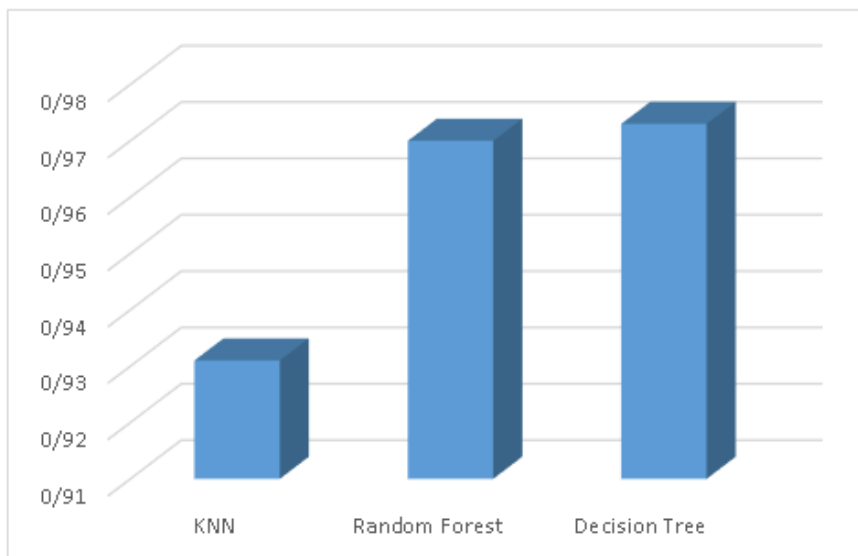
F1-Score	Recall	Precision	Accuracy	الگوریتم ها
0.963	0.925	0.931	0.966	KNN
0.978	0.957	0.97	0.98	Random Forest
0.979	0.966	0.973	0.996	Decision Tree

نتایج حاصل از ۳ معیار مطرح شده در خصوص دسته بندی در جدول فوق ارائه شده است که در ادامه این نتایج به صورت نمودار ارائه می گردد



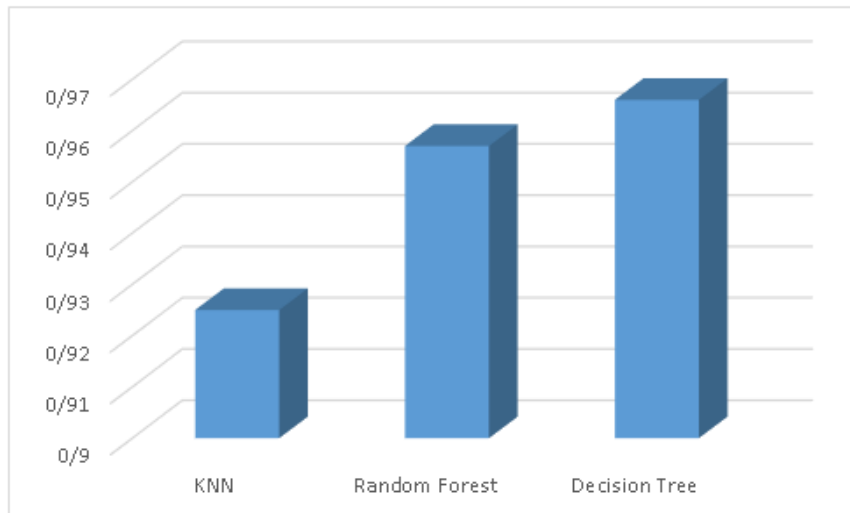
شکل ۲: مقایسه الگوریتم‌ها از نظر Accuracy

در شکل ۲ نشان داده می‌شود که الگوریتم درخت تصمیم دارای بیشترین میزان accuracy می‌باشد و مقدار accuracy نزدیک به ۱ را کسب کرده است.



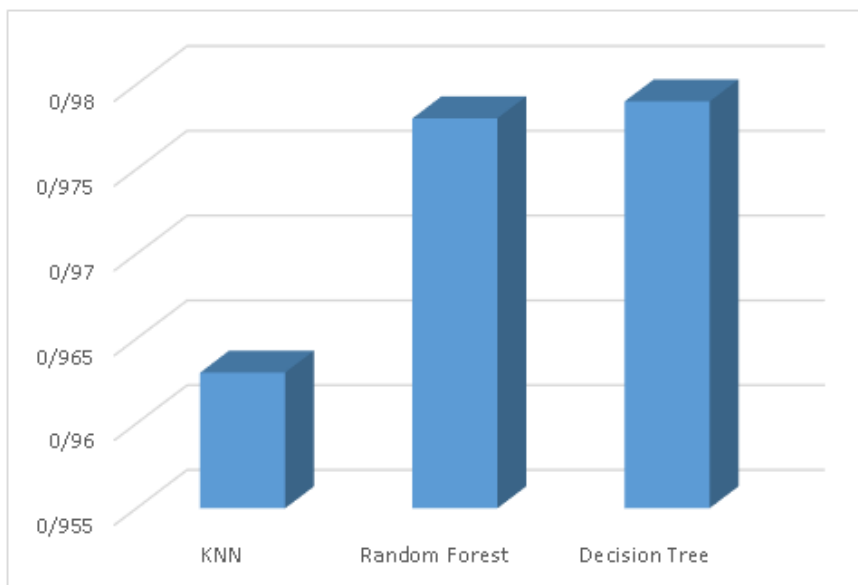
شکل ۳: مقایسه الگوریتم‌ها از نظر Precision

در شکل ۳ مشاهده می‌شود که درخت تصمیم همچنین نتایج بهتری نسبت به دو الگوریتم دیگر در خصوص Precision کسب کرده البته مشاهده می‌شود که جنگل تصادفی فاصله اندکی با درخت تصمیم دارد اما در مجموع درخت تصمیم از نظر precision دارای برتری می‌باشد.



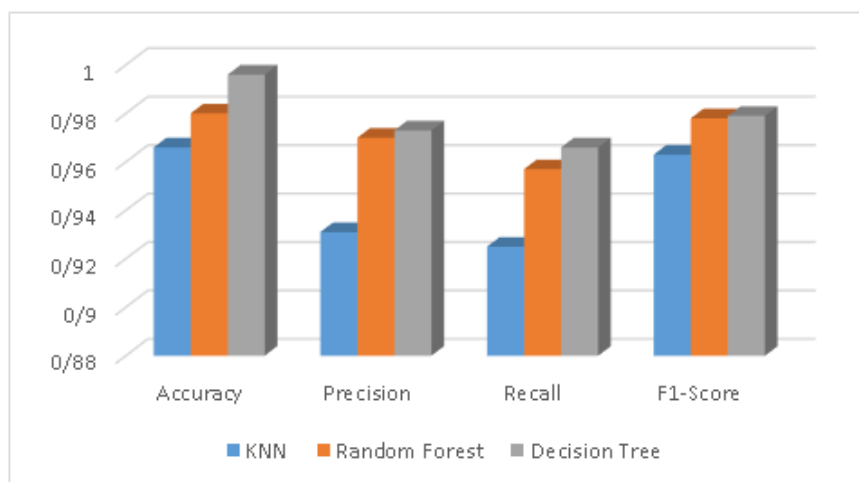
شکل ۴: مقایسه الگوریتم‌ها از نظر recall

در شکل ۴ مشاهده می‌شود الگوریتم درخت تصمیم نسبت به سایر الگوریتم‌ها از نظر recall نیز دارای برتری می‌باشد.



شکل ۵: مقایسه الگوریتم‌ها از نظر F1score

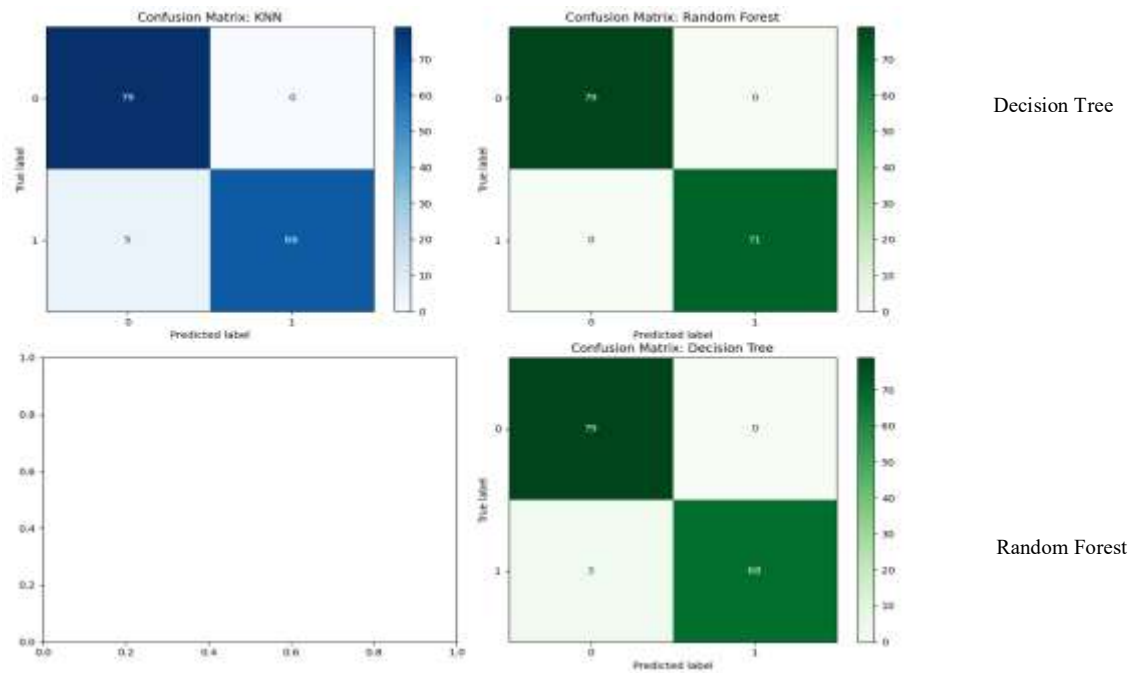
در شکل ۵ دیده می‌شود که الگوریتم درخت تصمیم با فاصله اندک نسبت به جنگل تصادفی از نظر معیار F1score قرار گرفته که نشانگر برتری این الگوریتم نسبت به دو الگوریتم مورد بررسی می‌باشد.



شکل ۶: مقایسه الگوریتم‌ها از نظر هر چهار معیار

در شکل ۶ تمامی الگوریتم‌ها در یک نگاه بررسی شده‌اند. نتایج نشان می‌دهد میله خاکستری که نشانگر نتایج حاصل از الگوریتم درخت تصمیم می‌باشد از سایر میله‌ها بلندتر می‌باشد و در نتیجه با توجه با اینکه بالا بودن مقادیر برای تمامی الگوریتم‌ها معیار برتری آن الگوریتم است در نتیجه الگوریتم درخت تصمیم الگوریتم برتر در تحقیق حاضر بوده و نتایج حاصل از آن معیار می‌باشد. الگوریتم کا نزدیکترین همسایه ضعیفترین الگوریتم در بین الگوریتم‌های موجود بوده و الگوریتم جنگل تصادفی در رتبه دوم قرار دارد البته باید تاکید شود که فاصله بین الگوریتم جنگل تصادفی و درخت تصمیم به خصوص در خصوص معیارهایی نظیر  $precision$  و  $f1score$  نسبت به سایر معیارها بسیار اندک می‌باشد.

در ادامه به منظور سنجش کارایی الگوریتم‌ها ماتریس در هم ریختگی یا کانفیوژن ارائه می‌شود که بر اساس اعداد روی قطر اصلی می‌توان برتری یک الگوریتم را تشکیل داد هر چه اعداد روی قطر اصلی در ماتریس سردرگمی بیشتر باشد می‌توان گفت الگوریتم مورد نظر دارای کارایی بیشتری می‌باشد نتایج در شکل ذیل ارائه گردیده است.

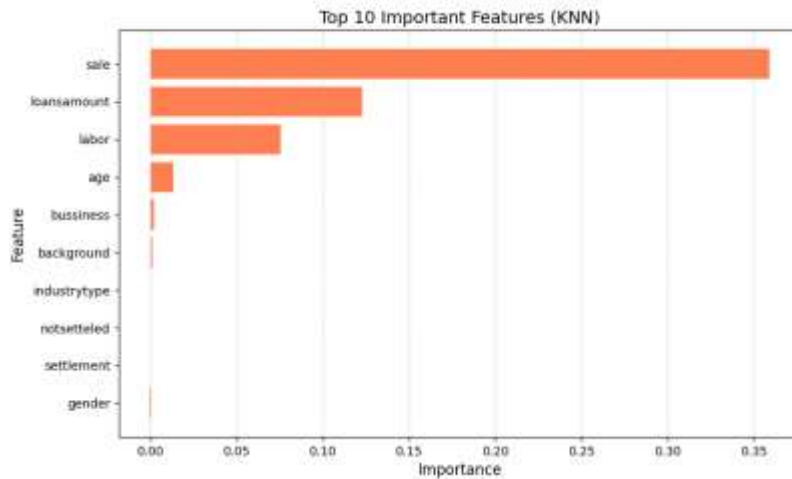


Decision Tree

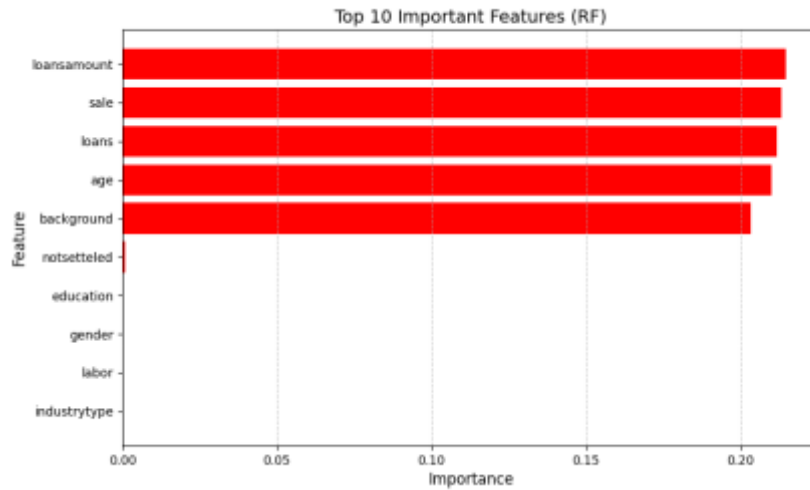
Random Forest

### شکل ۷: ماتریس در هم ریختگی

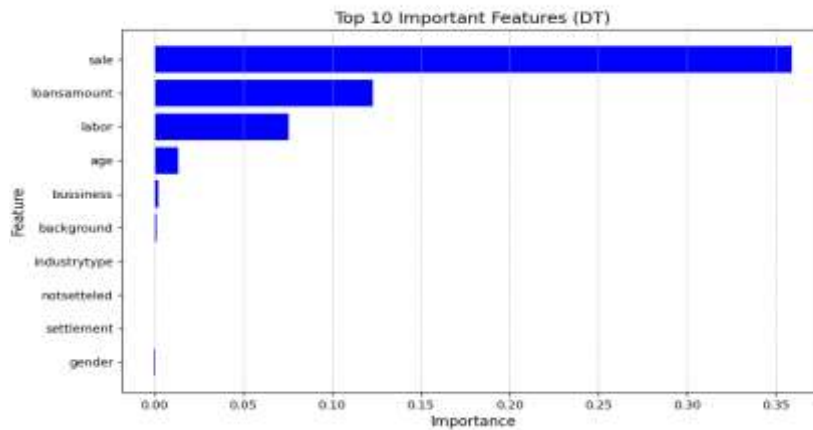
همانگونه که مشاهده می شود اعداد روی قطر اصلی برای الگوریتم درخت تصمیم بیشتر از سایر الگوریتم ها می باشد و با فاصله اندک می توان گفت الگوریتم درخت تصمیم از نظر ماتریس در هم ریختگی نیز برتر از سایر الگوریتم ها است. در ادامه به منظور تعیین اهمیت هر یک از متغیرها نمودار رتبه بندی مشخصه ها برای هر یک از الگوریتم ها ترسیم شده است که نتایج در ادامه ارائه می شود.



### شکل ۸: رتبه بندی ویژگیها بر اساس الگوریتم کا نزدیکترین همسایه



شکل ۹: رتبه بندی ویژگیها بر اساس الگوریتم جنگل تصادفی



شکل ۱۰: رتبه بندی ویژگیها بر اساس الگوریتم درخت تصمیم

همانگونه که مشاهده می شود در اشکال ۹ و ۱۰ و ۱۱ مقایسه ویژگیها بر اساس سه الگوریتم صورت گرفته که به دلیل آنکه الگوریتم درخت تصمیم نتایج دقیقتری را ارائه کرده و نشانگر کارایی بیشتری می باشد بنابراین نتایج حاصل از این الگوریتم در خصوص رتبه بندی مشخصه ها معیار قرار می گردد بر اساس نتایج حاصل از رتبه بندی ویژگیها به کمک الگوریتم درخت تصمیم می توان گفت میزان فروش در رتبه اول اثرگذاری به شمار می رود پس از آن میانگین وام دریافتی در سنوات گذشته دارای اثر بالایی می باشد. در رتبه سوم اهمیت تعداد نیروی کار تحت استخدام کسب و کار مورد نظر قرار گرفته و سپس سن مدیر کسب و کار یا فرد وام گیرنده قرار دارد.

رتبه چهارم مربوط به سابقه کسب و کار و رتبه پنجم از آن سابقه کار فرد می باشد که نشانگر بیشترین اثر در بازپرداخت وام و توانایی یا عدم توانایی در این خصوص می باشد. سایر ویژگیها که در واقع شش ویژگی بعدی می باشند در رتبه های بعدی اهمیت قرار داشته و اهمیت نسبتاً اندکی را نشان می دهند.

### نتیجه گیری

تحقیق حاضر به دنبال ارائه یک مدل هوشمند اعطای تسهیلات بانکی مبتنی بر کلان داده بر اساس مشخصه های جمعیت شناختی، وضعیت کسب و کار و پرداختهای گذشته مشتریان می باشد در این تحقیق تلاش شد متغیرهایی به مدل هوشمند ارائه تسهیلات اضافه شود که در تحقیقات پیشین کمتر مورد توجه قرار گرفته و لذا تحقیق حاضر به بسط و توسعه مدلهای اعطای تسهیلات هوشمند کمک نمود. نتایج نشان داد که مدل ارائه شده با ۱۲ مشخصه و در واقع متغیر ورودی می تواند تا ۹۹,۶ درصد بازپرداخت و ریسک آنرا تبیین نماید که نشانگر دقت بالایی می باشد در این تحقیق سه الگوریتم کا نزدیکترین همسایه، درخت تصمیم و جنگل تصادفی استفاده شد که هر سه الگوریتم دقت بالای ۹۵ درصد را نشان می دادند اما با فاصله اندک الگوریتم درخت تصمیم از نظر هر چهار معیار *f1 score, recall, accuracy, precision* دارای برتری بود ضمن اینکه نتایج رتبه بندی مشخصه ها نشان داد میزان فروش در رتبه اول اثرگذاری به شمار می رود پس از آن میانگین وام دریافتی در سنوات گذشته، تعداد نیروی کار تحت استخدام کسب و کار مورد نظر و سن وام گیرنده قرار دارد. سابقه کسب و کار و سابقه کار فرد می باشد در رتبه های بعدی اهمیت قرار دارند

بر اساس نتایج حاصل می توان گفت مدل ارائه شده مدل قوی با قدرت پیش بینی بالا در خصوص اعطای تسهیلات بانکی بوده و می تواند به عنوان یک مدل برای تصمیم گیری در خصوص اعطای تسهیلات بانکی مورد استفاده قرار گیرد نکته قابل توجه این است که مدل ارائه شده بر اساس ترکیبی از متغیرهای جمعیت شناختی، کسب و کار و سابقه پرداخت بوده و فراتر از صرفاً متغیرهای سابقه پرداخت عمل کرده که در تحقیقات پیشین مورد استفاده قرار گرفته است. یافته ها نشان می دهد از هر یک از دسته متغیرها، تعدادی نقش اساسی را ایفا می کند به گونه ای که فروش در خصوص وضعیت کسب و کار می باشد در حالیکه متغیر میانگین وام دریافتی در سنوات گذشته مربوط به سابقه پرداخت. تعداد نیروی کار تحت استخدام در ارتباط با متغیرهای وضعیت کسب و کار تعریف می

شود در حالیکه سن و سابقه کار وام گیرنده در بین متغیرهای جمعیت شناختی است. بنابراین می توان گفت تمامی زیر متغیرها به تفکیک هر دسته در پیش بینی بازپرداخت و تعیین ریسک پرداخت وام اثرگذار می باشند.

تحقیق آتی می تواند به توسعه مدل ارائه شده پرداخته و متغیرهای اصلی دیگری را به مدل افزوده و یا زیرمتغیرهای موجود در هر دسته بندی را افزایش داده و با استفاده از دیتاست مناسب سیستم بانکداری به ارزیابی مدل ارائه شده بپردازد.

### منابع

- Hotaa,L, Jaina,P, Kumar,A, A Comparative Performance Assessment for Prediction of Loan Approval in Financial Sector, *Procedia Computer Science* 258 (2025) 298–307.
- Thanh,C,N Huy,T,P, Hong,T,P & Quoc,A,N (2025) Credit risk prediction with corruption perception index: machine learning approaches, *Cogent Business & Management*, 12:1, 2461731, DOI: 10.1080/23311975.2025.2461731.
- Abdelhady ,S, Abdallah ,R, Fawzy ,A,M, Mohamed ,M,A, Salaheldin,Y A Data Driven Model for Predicting Loan Approval Using Machine Learning Approaches, *ERURJ* 2025, 4, 1, 2271-2289.
- Sriram,H, K Leveraging artificial intelligence and machine learning for next-generation credit risk assessment models, 2025 *EAJSE* Jan 2025, Volume 1, Issue 1.
- Fati,S,M A LOAN DEFAULT PREDICTION MODEL USING MACHINE LEARNING AND FEATURE ENGINEERING, *ICIC Express Letters ICIC International* c 2024 ISSN 1881-803X Volume 18, Number 1, January 2024 pp. 27–3.
- SINAP,V, A Comparative Study of Loan Approval Prediction Using Machine Learning Methods, *GU J Sci, Part C*, 12(2): 644-663 (2024).
- Dansana,D, Patro,G,K Mishra,B ,Prasad,V Razak,A, Anteneh Wogasso Wodajo, Analyzing the impact of loan features on bank loan prediction using Random Forest algorithm, *Engineering Reports*. 2024;6:e12707. [wileyonlinelibrary.com/journal/eng2](https://doi.org/10.1002/eng2.12707) 1 of 17 <https://doi.org/10.1002/eng2.12707>.
- Elrashidy,O, Bendary,M, Fouad,K, A. Elattar,M Intelligent Decision Support System for Loan Evaluation Using Machine Learning, DOI: 10.1109/MIUCC62295.2024.10783533.
- Hassan,A, A Machine Learning-Based Loan Eligibility Prediction System Utilizing Ensemble Technique with a Smart Application

- for Generating Installment Plans, *Journal of Computational Science and Applications (JCSA)*, Volume 1, Issue 1, 2024.
- Hrytsenko, L., Pakhnenko, O. Kuzior, A. & Kozhushko, I. (2024). Smart technologies in banking. *Financial Markets, Institutions and Risks*, 8(1), 81-93. [http://doi.org/10.61093/fmir.8\(1\).81-93.2024](http://doi.org/10.61093/fmir.8(1).81-93.2024).
- E. EL-GEMEIE ,Y, ABDELSALAM ,M, MOAWAD,I ENHANCING BANKING SERVICES THROUGH SMART DATA ANALYTICS FRAMEWORK, *Journal of Theoretical and Applied Information Technology* 15th October 2024. Vol.102. No. 19.
- Hussain,M,Z, Bank Loan Prediction System Using Machine Learning Models, 2024 IEEE 9th International Conference for Convergence in Technology (I2CT) Pune, India. Apr 5-7, 2024.
- Haque,F.M, BANK LOAN PREDICTION USING MACHINE LEARNING TECHNIQUES, *Journal of Theoretical and Applied Information Technology* 15th October 2024. Vol.104. No. 2.
- Raghavan,P, AI-Driven Data Management for Intelligent Loan Processing, *International Journal of All Research Education and Scientific Methods (IJARESM)*, ISSN: 2455-6211, Volume 12, Issue 6, June-2024.
- Hossain,S, Sajal,A, Jamee,S,S, Tisha,S,A, Siddique,T, Obaid,O, Karim Chy,S & Md Sayem Ul Haque. (2025). Comparative Analysis of Machine Learning Models for Credit Risk Prediction in Banking Systems. *The American Journal of Engineering and Technology*, 7(04), 22–33. <https://doi.org/10.37547/tajet/Volume07Issue04-04>.
- Zuama,R,Z, Ichsan,N, Pohan,A,B Azis,M,S Lase,M An implementation of machine learning on loan default prediction based on customer behavior, *Jurnal Info Sains : Informatika dan Sains* Volume 14 , Number 01, 2024, DOI 10.54209/infosains.v14i01 ESSN 2797-7889 (Online).
- Bi,S, Bao,W, Innovative Application of Artificial Intelligence Technology in Bank Credit Risk Management, *International Journal of Global Economics and Management* ISSN: 3005-9690 (Print), ISSN: 3005-8090 (Online) | Volume 2, Number 3, Year 2024.
- Chaudhry,R, Artificial Intelligence with Streamlining Payments and Lending for a Simpler Financial Ecosystem, 2024 International Conference on Emerging Smart Computing and Informatics (ESCI) AISSMS Institute of Information Technology, Pune, India. Mar 5-7, 2024.