



## Prediction the Choice of Financing for Start-ups using Machine Learning Algorithms and Behavioral Biases

Naime Niazi<sup>1</sup> and Hamideh Razavi<sup>2</sup>

1. MSc., Department of Industrial Engineering, Faculty of Engineering, University of Ferdowsi of Mashhad, Mashhad, Iran  
Email: [naimeniazi@gmail.com](mailto:naimeniazi@gmail.com)
2. Corresponding Author, Associate Prof. Department of Industrial Engineering, Faculty of Engineering, University of Ferdowsi of Mashhad, Mashhad, Iran, Email: [h-razavi@um.ac.ir](mailto:h-razavi@um.ac.ir)

### Article Info

**Article type:**  
Research Article

**Article history:**  
Received 17 Aug 2024  
Received in revised form 30 Aug 2024  
Accepted 15 Sep 2024  
Published online 21 Sep 2024

**Keywords:**  
Cluster boosting algorithm (Catboost),  
Cognitive Biase,  
Ensemble learning ,  
Financing,  
Startup.

### ABSTRACT

The aim of this paper is to predict financing methods to support decision-making for startup founders and their investors. Initially, factors influencing the choice of financing methods, including structural, demographic, and behavioral factors, were identified. These factors were then assessed using a questionnaire consisting of 32 items, which was sent online to startup founders. Based on 70 responses received and using algorithms including binary matching, classification chains, label power set, K-nearest neighbors, extreme gradient boosting, cluster boosting algorithm and random forest, the financing methods chosen by startups were predicted. Comparison of the results from the algorithms shows that the boosting ensemble algorithm, with an F1 score of 89 and precision of 85%, predicts the selected financing methods on the test dataset better than other algorithms. Additionally, data analysis indicates that startups are more inclined towards personal funding methods, which aligns with the prevalence of loss aversion bias among entrepreneurs. Following loss aversion, overconfidence, anchoring, and illusion of control biases were the most frequent among entrepreneurs.

**Cite this article:** Niazi, N., & Razavi, H. (2024). Prediction the Choice of Financing for Start-ups using Machine Learning Algorithms and Behavioral Biases. *Engineering Management and Soft Computing*, 10 (1), 238-261. DOI: <https://doi.org/10.22091/jemsc.2024.11203.1200>



© The Author(s)  
DOI: <https://doi.org/10.22091/jemsc.2024.11203.1200>

**Publisher:** University of Qom

## پیش‌بینی شیوه تامین مالی استارت‌آپ‌ها با استفاده از الگوریتم‌های یادگیری ماشین و در نظر گرفتن سوگیری‌های رفتاری

نعیمه نیازی<sup>۱</sup> و حمیده رضوی<sup>۲</sup>

۱. کارشناسی ارشد، مهندسی صنایع، دانشکده مهندسی دانشگاه فردوسی مشهد، مشهد، ایران. رایانامه: [naimeniazi@gmail.com](mailto:naimeniazi@gmail.com)

۲. نویسنده مسئول، دانشیار، مهندسی صنایع، دانشکده مهندسی دانشگاه فردوسی مشهد، مشهد، ایران رایانامه: [h-razavi@um.ac.ir](mailto:h-razavi@um.ac.ir)

اطلاعات مقاله	چکیده
<p>نوع مقاله: مقاله پژوهشی</p> <p>تاریخ دریافت: ۱۴۰۳/۰۵/۲۷</p> <p>تاریخ بازنگری: ۱۴۰۳/۰۵/۳۰</p> <p>تاریخ پذیرش: ۱۴۰۳/۰۶/۲۵</p> <p>تاریخ انتشار: ۱۴۰۳/۰۶/۳۱</p> <p>کلیدواژه‌ها: استارت‌آپ، الگوریتم تقویت دسته، تامین مالی، سوگیری‌های رفتاری، یادگیری گروهی.</p>	<p>هدف از این مقاله، پیش‌بینی شیوه‌های تامین مالی برای پشتیبانی تصمیم‌گیری بنیان‌گذاران استارت‌آپ‌ها و سرمایه‌گذاران آنهاست. در ابتدا عوامل موثر در انتخاب روش تامین مالی شامل عوامل ساختاری، جمعیت‌شناختی و رفتاری شناسایی شدند. سپس این عوامل با استفاده از پرسشنامه‌ای مشتمل بر ۳۲ گویه، به صورت آنلاین برای بنیان‌گذاران استارت‌آپ‌ها ارسال گردید. براساس ۷۰ پاسخ دریافت‌شده و استفاده از الگوریتم‌های تطابق دودویی، زنجیره‌های طبقه‌بندی، مجموعه توان برجسب‌ها، K نزدیک‌ترین همسایه، تقویت گرادیان شدید، تقویت دسته و جنگل تصادفی به پیش‌بینی شیوه‌های تامین مالی که استارت‌آپ‌ها انتخاب می‌کنند، پرداخته شد. مقایسه نتایج حاصل از الگوریتم‌ها نشان می‌دهد که الگوریتم تقویت دسته با شاخص ارزیابی F1 معادل ۸۹ و دقت معادل ۸۵ درصد بهتر از سایر الگوریتم‌ها، روش‌های تامین مالی منتخب را بر روی مجموعه داده آزمون پیش‌بینی می‌نماید. همچنین تحلیل داده‌ها نشان می‌دهد که استارت‌آپ‌ها بیشتر به روش پرداخت شخصی برای تامین مالی تمایل دارند که با فراوانی سوگیری زیان‌گریزی در میان کارآفرینان همسو است. همچنین پس از سوگیری زیان‌گریزی، سوگیری‌های اعتمادبه‌نفس بیش‌ازحد، لنگرانداختن و توهم کنترل، بیشترین فراوانی را در میان کارآفرینان داشتند.</p>

**استناد:** نیازی، نعیمه؛ رضوی، حمیده. (۱۴۰۳). «پیش‌بینی شیوه تامین مالی استارت‌آپ‌ها با استفاده از الگوریتم‌های یادگیری ماشین و در نظر گرفتن سوگیری‌های رفتاری». *مدیریت مهندسی و رایانش نرم*، دوره ۱۰ (۱)، صص: ۲۳۸-۲۶۱. <https://doi.org/10.22091/jemsc.2024.11203.1200>



## ۱) مقدمه

استارت‌آپ‌ها، از شناسایی فرصت توسط کارآفرینان شکل می‌گیرند. گروهی از افراد باهم کار می‌کنند تا محصولی را بسازند یا خدمتی را به شکلی متفاوت از آنچه پیش از آن وجود داشته، به مشتریان ارائه کنند (Baldridge, 2022). در این کسب‌وکارهای ساختاریافته، نقش کارآفرینان در تصمیم‌گیری‌ها، پرننگ‌تر از شرکت‌های بزرگ و ساختاریافته است (Nguyen & Canh, 2021). یکی از این چالش‌ها و مسائل محدودکننده که از ابتدای شکل‌گیری تا بلوغ یک کسب‌وکار وجود دارد، مسئله انتخاب شیوه تامین مالی است.

در ابتدای کار، فعالیت‌ها برای سرمایه‌گذاران، دارای ابهامات زیادی در حوزه‌های مختلف از جمله زمینه فعالیت، توانایی سرمایه‌گذاری و جریانات نقدی است که این مسئله برآورد ریسک سرمایه‌گذاران از سرمایه‌گذاری را افزایش می‌دهد (Bauer et al., 2023). بنابراین کسب‌وکارها در هنگام شکل‌گیری به تنهایی، ناچار به تامین منابع مالی موردنیاز هستند. مقدار سرمایه قابل دسترس از این طریق، به استثنای کسب‌وکارهایی که بنیان‌گذاران ثروتمندی دارند، محدود است. بنابراین کارآفرینان تازه کاری که به سرمایه‌گذاری اولیه بیشتری نیاز دارند، به دنبال منابع مالی دیگری مانند وام‌های بانکی، فرشتگان کسب‌وکار و یا سرمایه‌گذاران خطرپذیر هستند. تامین مالی از طریق این منابع با توجه به عدم تقارن اطلاعاتی میان استارت‌آپ‌ها و سرمایه‌گذاران، اغلب پرهزینه است (Honjo et al., 2022). در صورت موفقیت در تامین سرمایه موردنیاز از طریق هریک از روش‌های تامین مالی غیرشخصی، بدهی‌هایی برای استارت‌آپ‌ها ایجاد می‌شود که این بدهی‌ها نرخ شکست کسب‌وکارها را در ابتدای کار افزایش می‌دهند (Berger & Hottenrott, 2021).

اغلب تصمیم‌گیری‌های مالی در محیط کسب‌وکارهای پیچیده و متغیر صورت می‌گیرد که با عدم قطعیت زیادی همراه هستند (Hallo et al., 2020)، این عدم قطعیت‌ها و پیچیدگی‌ها، موجب تشدید اولویت‌شهود کارآفرینان بر عقلانی‌تری که نظریه‌های متعارف اقتصاد و مالی بر آن استوار است، می‌گردد. هدف از آشنایی با سوگیری‌های رفتاری در فرآیند تصمیم‌گیری، شناسایی موقعیت‌هایی است که در آنها این سوگیری‌ها می‌توانند خود را بروز دهند تا بدینوسیله افراد بتوانند از اشتباهات در تصمیم‌گیری، به خصوص تصمیم‌گیری‌های مالی اجتناب کنند (Lien et al., 2022). تاکنون تعداد زیادی سوگیری رفتاری شناسایی شده و روزبه‌روز بر این تعداد افزوده می‌شود (Harris, 2021).

این پژوهش به دنبال پیش‌بینی انتخاب شیوه تامین مالی در استارت‌آپ‌ها با در نظر گرفتن ویژگی‌های ساختاری، جمعیت‌شناختی و رفتاری است. برای رسیدن به این هدف، از پرسشنامه برای جمع‌آوری اطلاعات استفاده می‌شود و از آنجاییکه نظریات ساختاریافته‌ای در این زمینه برای حل مسئله وجود ندارد از ابزارهای داده‌محور برای تبیین آن استفاده می‌شود (Simon et al., 2000). منظور از عوامل ساختاری: تعداد کارمندان، سرمایه اولیه، درآمد، ماهیت محصول و امثال آن است. عوامل جمعیت‌شناختی عبارتند از: سن، جنسیت، تحصیلات و درآمد و چهار سوگیری شناختی شامل اعتمادبه‌نفس بیش‌ازحد، توهم کنترل، زیان‌گریزی و لنگرانداختن است.

در این مقاله، در ابتدا خلاصه‌ای از تحقیقات پیشین به تفکیک در زمینه عوامل ساختاری استارت‌آپ‌ها، عوامل جمعیتی و رفتاری کارآفرینان که در این پژوهش مورد بررسی قرار گرفته‌اند و تصمیم‌گیری انسانی توضیح داده می‌شود. سپس

مدل مسئله، پردازش داده‌ها و الگوریتم‌های استفاده‌شده به تفکیک معرفی شده، نتایج حاصل از اجرای الگوریتم‌ها مورد مقایسه قرار گرفته و در انتها الگوریتم برتر شناسایی می‌شود. به کمک این روش، می‌توان برای هر استارت‌آپ جدید، براساس اطلاعات اعضای آن و شناسایی سوگیری‌های رفتاری ایشان، روش تامین مالی مناسب را پیش‌بینی و یا توصیه نمود.

## ۲) پیشینه پژوهش

استارت‌آپ‌ها به‌عنوان سازمان‌های موقتی تعریف می‌شوند که برای جست‌وجوی مدل کسب‌وکاری قابل تکرار و مقیاس‌پذیر طراحی شده‌اند (Blank, 2010). این کسب‌وکارها اغلب سرمایه راه‌اندازی پایینی دارند و به دلایلی چون ریسک شکست بالا، وضعیت اعتباری کارآفرینان و هزینه‌های بالای مبادله، دسترسی کمی به روش‌های تامین مالی بیرونی دارند. این مسئله، شیوه انتخاب تامین مالی برای این کسب‌وکارها را در بستری متفاوت از سایر شرکت‌ها قرار می‌دهد (Bollaert et al., 2021). روش‌های تامین مالی رایج در ادبیات برای تامین بودجه موردنیاز این کسب‌وکارهای نوپا شامل ۱۰ روش تامین مالی: ۱. پرداخت شخصی ۲. کمک خانواده، دوستان و نزدیکان ۳. وام‌های بانکی ۴. وام‌های همتابه‌همتا ۵. تامین مالی جمعی ۶. سرمایه‌گذاران فرشته ۷. سرمایه‌گذاری شتاب‌دهنده‌ها ۸. سرمایه‌گذاری خطرپذیر ۹. سرمایه‌گذاری خطرپذیر شرکتی ۱۰. عرضه اولیه به‌همراه هستند.

اولین و دردسترس‌ترین این روش‌ها، پرداخت شخصی است. در این روش، کارآفرینان برای سرمایه‌گذاری از پول و سرمایه شخصی خود استفاده می‌کنند. در ادبیات کارآفرینی، عبارت خودراه‌اندازی<sup>۱</sup> رایج‌تر از تامین مالی شخصی است. خودراه‌اندازی علاوه بر تامین مالی از طریق سرمایه شخصی شامل یافتن راه‌های خلاقانه برای مدیریت و اداره استارت‌آپ به‌گونه‌ای است که کمترین نیاز به سرمایه را پیدا کند (Tech, 2018).

روش بعدی، تامین مالی از طریق خانواده، دوستان و نزدیکان است. باید توجه داشت که این روش می‌تواند در روابط شخصی افراد مشکلاتی را ایجاد کند و لازم است تا طرفین پیش از سرمایه‌گذاری تبعات آن مانند احتمال شکست، سهم سرمایه‌گذار و زمان خروج را در نظر بگیرند (Hill, 2018).

وام‌های بانکی یکی دیگر از روش‌های تامین مالی موجود هستند. برای استارت‌آپ‌ها به‌خصوص در مراحل اولیه که کارآفرینان اعتبار و دارایی‌های ملموسی ندارند، تامین مالی از طریق وام‌های بانکی دشوار است چراکه وام‌دهندگان به این دلیل که به دنبال کاهش ریسک هستند، عرضه را حتی در صورت تقاضا با نرخ‌های کارمزد بالاتر، محدود می‌کنند (Dhewanto et al., 2019).

وام‌های همتابه‌همتا نیز از جدیدترین ابزارهای تامین مالی هستند. در این نوع وام‌ها، وام‌گیرندگان اطلاعات و شرایط خود را اعلام می‌کنند و وام‌دهندگان تصمیم می‌گیرند که در وام مشارکت کنند یا خیر (Basha et al., 2021). از دیگر روش‌های تامین مالی جدید و روبه‌رشد، تامین مالی جمعی است (Cai et al., 2021). در سال‌های اخیر، نهادهای تامین مالی جمعی در ایران به شیوه وام شکل گرفته و با رویدادهایی تحت عنوان سکوی سرمایه‌گذاری به جذب سرمایه موردنیاز متقاضیان می‌پردازند.

<sup>۱</sup> Bootstrapping

از دیگر روش‌های تامین مالی، سرمایه‌گذاران فرشته هستند. این افراد سرمایه اولیه استارت‌آپ‌ها و کسب‌وکارهای نوپا را با استفاده از سرمایه شخصی تامین می‌کنند. سرمایه‌گذاری فرشتگان در شرایط عدم اطمینان بالایی شکل می‌گیرد و در این وضعیت برداشت‌های فردی فرشتگان در ارزیابی و تفسیر کسب‌وکارها نقش پررنگی در فرآیند ارزیابی و سرمایه‌گذاری آنان ایفا می‌کند (Lien et al., 2022).

شتاب‌دهنده‌ها در بسیاری از پژوهش‌ها به‌عنوان یک روش تامین مالی در نظر گرفته می‌شوند اما بطور کلی علاوه بر تامین سرمایه و فضای کار به کسب‌وکارهای نوپا کمک می‌کنند تا محصولات اولیه خود را تولید کنند و همچنین فرصت‌های ارتباطی برای آنان ایجاد می‌کنند (Harris, 2021).

روش تامین مالی بعدی، سرمایه‌گذاران خطرپذیر<sup>۲</sup> است. سرمایه‌گذاران خطرپذیر در واقع واسطه‌های مالی هستند که در شرکت‌های خصوصی سرمایه‌گذاری می‌کنند و در طی سرمایه‌گذاری نقش پررنگی در نظارت بر هزینه‌کرد مبلغ سرمایه‌گذاری دارند. هدف این سرمایه‌گذاران کمک به تامین مالی کسب‌وکارهای نوپا و خروج پس از کسب بازده موردانتظار سرمایه‌گذاری است (Metrick & Yasuda, 2021).

سرمایه‌گذاران خطرپذیر شرکتی<sup>۳</sup>، روش تامین مالی بعدی هستند. این سرمایه‌گذاران شرکت‌های بزرگی هستند که به سرمایه‌گذاری در کسب‌وکارهایی می‌پردازند که بیشترین شباهت را با شرکت‌هایشان داشته باشند و بدینوسیله امکان رشد شرکت‌های استارت‌آپی مشابه که دارای دانش فنی تکمیل‌کننده و یا مشتریان جدیدی هستند را فراهم می‌کنند (Hegeman & Sørheim, 2021).

عرضه اولیه به‌امهر<sup>۴</sup>، روش تامین مالی جدیدتری نسبت به سایر روش‌ها در ادبیات موضوع است. این روش شبیه تامین مالی جمعی از طریق سهام است. در این روش استارت‌آپ‌ها برای تامین سرمایه اولیه، توکن که اغلب این توکن‌ها یک ارز دیجیتال هستند را با تخفیف و قیمتی پایین‌تر از قیمت عادی برای فروش با ارزهای دیجیتال مرسوم مانند بیت‌کوین و اتریوم عرضه می‌کنند (Fisch, 2019).

پس از شناسایی روش‌های مرسوم تامین مالی استارت‌آپ‌ها، لازم است مهمترین عوامل موثر بر تصمیم‌گیری انتخاب شیوه تامین مالی شناسایی شوند. در امور مالی کارآفرینی، تصمیم‌گیری از دو جنبه عدم تقارن اطلاعاتی میان سرمایه‌گذاران و کارآفرینان و عدم دسترسی به عوامل تسهیل‌گر نظیر سرمایه اجتماعی، با شرکت‌های دیگر متفاوت است (Ding et al., 2015; Fernandez, 2021). انتظار می‌رود که کارآفرینان به دلیل قرار گرفتن در شرایط تصمیم‌گیری با عدم قطعیت زیاد، محدودیت زمانی و اطلاعات نامتقارن نسبت به سایرین، از رفتار منطقی و عقلایی فاصله بگیرند و در تصمیمات دچار سوگیری‌های متعددی شوند (Metrick & Yasuda, 2021). سوگیری‌های رفتاری یکی از چندین عاملی هستند که بر تصمیم‌گیری مالی کارآفرینان اثر می‌گذارند. علاوه بر سوگیری‌ها، عوامل متعدد دیگری شامل عوامل ساختاری و جمعیتی نیز بر انتخاب روش تامین مالی یا ساختار سرمایه شرکت‌ها تاثیرگذار هستند.

<sup>2</sup> Venture capital (VC)

<sup>3</sup> corporate venture capital (CVC)

<sup>4</sup> Initial Coin Offering (ICO)

## ۲-۱) عوامل ساختاری استارت‌آپ‌ها

عوامل ساختاری متفاوت، بسترهای تصمیم‌گیری متفاوت و انتخاب‌های متمایزی را برای استارت‌آپ‌ها فراهم کرده که این امر موجب پیامدهای متفاوتی برای کسب و کارها خواهد شد. از بین عوامل ساختاری که در ادبیات موضوع مورد بررسی قرار گرفته‌است، با توجه به پژوهش‌های پیشین و امکان دسترسی به اطلاعات، برخی از عوامل که در ادامه تشریح می‌شوند برای بررسی در این پژوهش انتخاب شدند.

۱. **عمر استارت‌آپ:** برخی پژوهش‌ها نشان می‌دهند که اگرچه شرکت‌های جوان نرخ شکست بالایی دارند اما مشروط به بقا، نرخ رشد بالاتری را در مقایسه با شرکت‌های قدیمی‌تر نشان می‌دهند (Ouimet & Zarutskie, 2014).
۲. **اندازه استارت‌آپ:** شرکت‌های کوچکتر مزایا و انعطاف‌پذیری بیشتری در محیط بازارهای در حال تغییر دارند و بنابراین احتمالاً بیشتر از شرکت‌های بزرگتر زنده می‌مانند (Lybaert & Umans, 2022).
۳. **ثبات اختراع:** وقتی شرکت‌ها در تامین مالی اولیه و بدون سابقه تصمیم به جذب منابع مالی خارجی می‌گیرند، باید راه‌هایی برای نشان دادن کیفیت خود به سرمایه‌گذاران بالقوه پیدا کنند. ثبت اختراع‌ها می‌تواند به عنوان عاملی برای کیفیت در نظر گرفته شود (Vo, 2019).

## ۲-۲) عوامل جمعیت‌رفتاری کارآفرینان

کارآفرین به عنوان فردی که برای تبدیل ایده‌ها به فرصت‌های تجاری و ایجاد ارزش فکر و عمل می‌کند، شناخته می‌شود (Tech, 2018). در این پژوهش، واژه‌های کارآفرین و بنیان‌گذار استارت‌آپ در یک معنا بکار برده می‌شوند و عوامل جمعیت‌شناختی مانند سن، جنسیت و تحصیلات بر تصمیم‌گیری تامین مالی استارت‌آپ‌ها مورد بررسی قرار می‌گیرد. پژوهش‌ها نشان می‌دهند که شرکت‌های رشد‌محور و افرادی که آنها را راه‌اندازی می‌کنند، این دیدگاه رایج را تکرار می‌کنند که موفق‌ترین و متحول‌کننده‌ترین سرمایه‌گذاری‌های جدید از سوی جوانان می‌آیند (Azoulay et al., 2020). بنابراین سن کارآفرینان یک عامل موثر در تصمیم‌گیری است. عامل موثر بعدی جنسیت است. پژوهش‌های مختلف نشان می‌دهند که کارآفرینان زن تقریباً بطور میانگین ۳۰ درصد از جمعیت بنیان‌گذاران استارت‌آپ‌ها را در طول زمان و در کشورهای مختلف تشکیل می‌دهند و از این تعداد ۱۰ تا ۱۵ درصد از آنها موفق به دریافت سرمایه خصوصی و سرمایه خطرپذیر می‌شوند (Hebert, 2020). سوابق تحصیلی بنیان‌گذاران استارت‌آپ‌ها می‌تواند به عنوان عاملی موثر و محرکی قوی برای عملکرد آنان تلقی شود (Franco et al., 2021).

## ۲-۳) عوامل رفتاری کارآفرینان

پژوهش بر روی تاثیر سوگیری‌های رفتاری با هدف درک این مسئله که کارآفرینان چگونه آگاهانه یا غیرآگاهانه تصمیم می‌گیرند، از اواسط دهه ۱۹۹۰ آغاز شده‌است و از آن زمان تاکنون به یکی از حوزه‌های مهم پژوهش‌های کارآفرینی تبدیل شده‌است (Rosyidah & Pratikto, 2022). جدول (۱) مرور مختصری است بر پژوهش‌هایی که در این زمینه مورد بررسی قرار گرفتند.

## جدول ۱۰. پژوهش‌های گذشته در مورد تاثیر سوگیری‌های مختلف

ردیف	نام نویسندگان	سال انتشار	سوگیری‌های مورد بررسی	روش پژوهش
۱	ژانگ و همکاران <sup>۵</sup> ( Zhang & Cueto, 2017 )	۲۰۱۸	اعتماد به نفس بیش‌ازحد <sup>۶</sup> ، خوش‌بینی <sup>۷</sup> ، خوداستنادی <sup>۸</sup> ، توهم کنترل <sup>۹</sup> ، قانون اعداد کوچک <sup>۱۰</sup> ، مغالطه در برنامه‌ریزی <sup>۱۱</sup> و تشدید تعهد <sup>۱۲</sup>	مروری
۲	زهیرا <sup>۱۳</sup> و همکاران ( Zahera & Bansal, 2018 )	۲۰۱۸	اعتماد بیش‌ازحد، رفتار گله‌ای، حسابداری ذهنی <sup>۱۴</sup> ، تایید <sup>۱۵</sup> ، واپس‌نگری <sup>۱۶</sup> ، نمادگرایی، اثر اهدا <sup>۱۷</sup> یا بیزاری از واگذاری، زیان‌گریزی، قالب‌بندی <sup>۱۸</sup> ، خانه <sup>۱۹</sup> ، خود اسنادی، محافظه‌کاری <sup>۲۰</sup> ، پشیمانی‌گریزی، لنگرانداختن، تازگی <sup>۲۱</sup> و اثر وضع <sup>۲۲</sup>	مروری
۳	رامالکشمی و همکاران <sup>۲۳</sup> ( Ramalakshmi et al., 2019 )	۲۰۱۹	اعتماد به نفس بیش‌ازحد، پشیمانی‌گریزی، رفتار گله‌ای و نمادگرایی	معادلات ساختاری
۴	دومینیک و همکاران <sup>۲۴</sup> ( Dominic & Gupta, 2020 )	۲۰۲۰	سوگیری اعتماد به نفس بیش‌ازحد، مغالطه قمارباز، لنگرانداختن و رفتار گله‌ای	آزمون‌های فرض با استفاده از همبستگی و رگرسیون
۵	آبدین و همکاران <sup>۲۵</sup> ( Ul Abdin et al., 2022 )	۲۰۲۲	خوداستنادی <sup>۲۶</sup> ، خوش‌بینی و توهم کنترل	کیفی و کمی

پس از بررسی پژوهش‌های پیشین و با توجه به نظر خبرگان در این پژوهش، چهار سوگیری اعتماد به نفس بیش‌ازحد، زیان‌گریزی، لنگرانداختن و توهم کنترل بر تصمیم‌گیری تامین مالی در بین بنیان‌گذاران استارت‌آپ‌ها انتخاب شدند که در ادامه به تفکیک شرح داده می‌شوند.

<sup>5</sup> Zhang et al  
<sup>6</sup> Overconfidence  
<sup>7</sup> Overoptimism  
<sup>8</sup> Self-serving attribution  
<sup>9</sup> Illusion of control  
<sup>10</sup> The law of small numbers  
<sup>11</sup> Planning fallacy  
<sup>12</sup> Escalation of commitment  
<sup>13</sup> Zahera et al  
<sup>14</sup> Mental Accounting  
<sup>15</sup> Confirmation  
<sup>16</sup> Hindsight  
<sup>17</sup> Endowment Effect  
<sup>18</sup> Framing  
<sup>19</sup> Home  
<sup>20</sup> Conservatism  
<sup>21</sup> Recency  
<sup>22</sup> Disposition  
<sup>23</sup> Ramalakshmi et al.  
<sup>24</sup> Dominic et al  
<sup>25</sup> Abdin et al  
<sup>26</sup> Self-attribution bias

اعتماد به نفس بیش از حد یکی از سوگیری‌های رفتاری رایج است. این سوگیری به معنای تمایل افراد به دست بالا گرفتن برآوردهایشان است (Bazerman & Moore, 2012). اعتماد به نفس بیش از حد موجب ناچیز انگاشتن ریسک در بسیاری از تصمیم‌های کارآفرینان می‌شود (Simon & Houghton, 2003).

در سوگیری زیان‌گریزی، اغلب باخت‌ها بسیار بزرگتر از بردها به نظر می‌رسند و افراد از باخت یا زیان، گریزان هستند. به بیان دیگر برخلاف نظریه‌های کلاسیک اقتصادی افراد بجای آنکه به دنبال حداکثر کردن مطلوبیت موردانتظار<sup>۲۷</sup> باشند، احتمالاً در جستجوی راهی هستند که از زیان جلوگیری کنند و بدیهی است که در معرض سوگیری زیان‌گریزی باشند (Kahneman, 2011).

در تصمیم‌گیری‌های اقتصادی اثر لنگرانداختن بسیار قوی و مهم به نظر می‌رسد. این سوگیری زمانی رخ می‌دهد که افراد از قبل، تخمین کمیتی برای آن ارزش خاص قائل باشند (Kahneman, 2011). سوگیری توهم کنترل، زمانی رخ می‌دهد که افراد باور دارند می‌توانند بر اتفاقات و عواملی که در کنترل آنها نبوده و دیگر عوامل محیطی در وقوع آنان نقش مهمی را ایفا می‌کنند، کنترل داشته باشند (Langer, 1975).

## ۲-۴) پیش‌بینی تصمیم‌گیری‌های انسانی

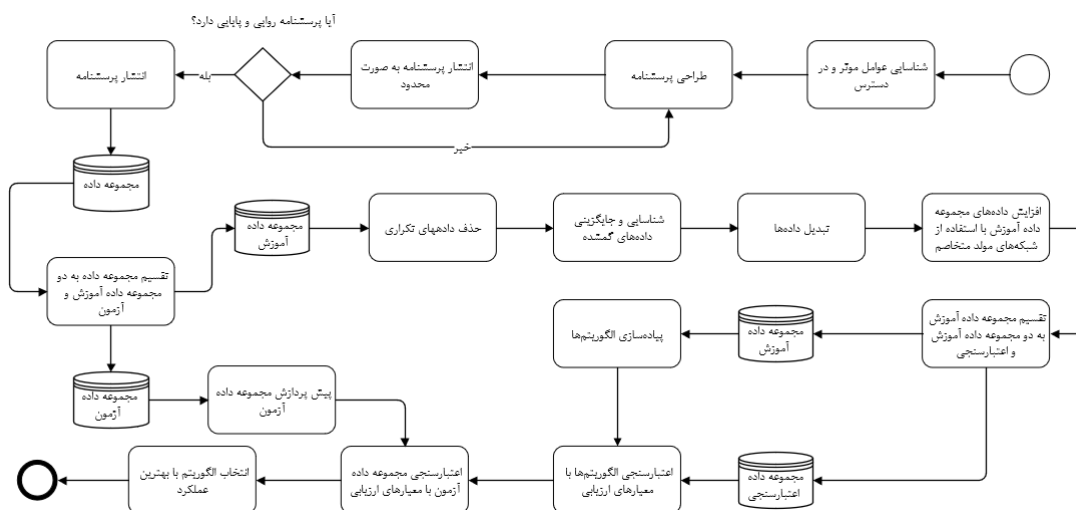
پژوهش‌ها نشان می‌دهند که ثبت جزئیات موثر بر تصمیم‌گیری انسانی و پیش‌بینی یک تصمیم و عمل براساس آن پیش‌بینی، مسئله ساده‌ای نیست و چالش‌های زیادی را به همراه دارد. برخی از پژوهش‌ها، به سه رویکرد برای پیش‌بینی تصمیم‌گیری‌های انسانی اشاره می‌کنند که این رویکردها عبارتند از تجربه‌محور، داده‌محور و ترکیبی. مدل‌های تجربه‌محور برای پیش‌بینی تصمیم‌گیری انسانی شامل تئوری‌ها و نظریه‌هایی هستند که توسط متخصصان هر حوزه بیان می‌شوند. یک مدل مبتنی بر داده، عموماً بر مبنای الگوریتم‌های یادگیری ماشین است که با استفاده از مجموعه داده‌های مختلف درباره تصمیم‌های افراد آموزش داده می‌شود. اگرچه مشاهده شده‌است که این رویکرد، از نظر دقت در پیش‌بینی انتخاب بسیار عالی است اما در ارائه بینش و استدلال در مورد چرایی ترجیحات، عملکرد مناسبی ندارد. در رویکردهای ترکیبی، هر دو رویکرد تجربه‌محور و داده‌محور ترکیب می‌شوند. در مورد اینکه برای یک مسئله مفروض، کدامیک از رویکردهای پیش‌بینی ذکر شده، اتخاذ شود به عوامل زیادی بستگی دارد. اگرچه بطور شهودی بیان می‌شود که هیچ الگوریتم واحدی بهترین کارآیی را در همه مسائل و حوزه‌ها ندارد اما با توجه به عواملی مانند در دسترس بودن داده‌ها، دانش تخصصی و ویژگی‌های حوزه مورد پژوهش، یکی از این رویکردها انتخاب می‌شوند (Rosenfeld & Kraus, 2018). براساس تعاریف فوق و با توجه به ماهیت مسئله، در این پژوهش از رویکرد داده‌محور استفاده شده‌است.

## ۳) روش‌شناسی پژوهش

شکل (۱) مراحل این پژوهش برای مدل‌سازی انتخاب شیوه تامین مالی استارت‌آپ‌ها را به اختصار نشان می‌دهد.

<sup>27</sup> Expected Utility





شکل ۲۶. مراحل پژوهش

### ۳-۱) جمع‌آوری مجموعه‌داده

شیوه جمع‌آوری داده در این پژوهش پرسشنامه است. پیش از استفاده از پرسشنامه باید از دقت و سازگاری این ابزار اطمینان حاصل پیدا کرد. دقت و سازگاری پرسشنامه‌ها در ادبیات موضوع، تحت عناوین روایی و پایایی شناخته می‌شوند. معیار روایی تضمین‌کننده این است که آیا نتایج، بازتابی واقعی از ساختار مورد مطالعه هستند؟ و آیا معیار پایایی، معیاری برای تعیین سازگاری و تکرارپذیری نتایج یک پژوهش است و تضمین می‌کند که نتایج به دلیل شانس یا خطای تصادفی به دست نیامده‌اند. در این پژوهش و با توجه به ماهیت مسئله برای پایایی از روش آزمون-آزمون مجدد که عبارت است از تکرار استفاده از یک ابزار اندازه‌گیری یکسان مانند پرسشنامه برای افراد یکسان تحت شرایط یکسان و پس از مدت‌زمان مشخص، استفاده می‌گردد. در این آزمون، اگر ضریب همبستگی پاسخ‌های آزمون و آزمون مجدد بزرگتر یا مساوی ۰/۷ باشد، پرسشنامه از نظر پایایی قابل‌اتکا و معتبر در نظر گرفته می‌شود. به‌منظور بررسی پایایی، پرسشنامه در دو نوبت به فاصله زمانی ۱۴ روز برای ۱۰ نفر ارسال شد و ضریب همبستگی پاسخ‌های هر دو پرسشنامه، بزرگتر از ۹۴ درصد به‌دست آمد. برای روایی نیز از روایی صوری در این پژوهش استفاده می‌شود. این معیار، درجه‌ای که سوالات آنچه را که قرار است اندازه‌گیری کنند، اندازه‌گیری می‌کنند یا خیر را می‌سنجد. این معیار توسط کارشناسانی که محتوای پرسشنامه را ارزیابی می‌کنند، تعیین می‌شود (Bolarinwa, 2015).

برای جمع‌آوری داده‌ها باید اطمینان حاصل کرد که داده‌ها تنها محدود به منطقه یا صنعت خاصی نباشند و همچنین نتایج تنها ناشی از تجزیه و تحلیل بازماندگان یا برندگان نباشد (Cassar, 2004). در همین راستا در هنگام جمع‌آوری داده به سوگیری بقا<sup>۲۸</sup> یا بازمانده توجه شد (Elston, 2021). برای تعیین حداقل حجم نمونه با توجه به اینکه مسئله مورد بررسی دارای حجم جامعه نامعین است، از رابطه ۱ که تحت عنوان رابطه کوکران<sup>۲۹</sup> شناخته می‌شود، استفاده می‌گردد. با استفاده از این رابطه، تعداد حداقل نمونه موردنیاز از طریق برآورد واریانس نمونه اولیه و درصد خطا محاسبه می‌شود (Israel, 1992)

<sup>28</sup> Survivorship Bias

<sup>29</sup> Cochran

$$n = \frac{Z_{\alpha}^2 \times p(1-p)}{d^2} \quad \text{رابطه ۱}$$

حداقل نمونه موردنیاز با استفاده از رابطه ۱، با انحراف معیار ۰/۳۶ داده‌های اولیه و با در نظر گرفتن ۸ درصد خطا، ۷۷ نمونه به دست آمد.

لینک پرسشنامه در بازه زمانی برای بیش از ۳۰۰ بنیان‌گذار استارت‌آپ از طریق بسترهای مختلف فضای مجازی ارسال و مجموعاً ۷۰ پاسخ دریافت و با توجه به اختلاف نزدیک با حداقل تعداد نمونه داده‌های حاصل از آن جهت مدل‌سازی استخراج گردید.

### ۳-۲) آماده‌سازی داده‌ها

مجموعه داده‌ها اغلب دارای مقادیر گمشده هستند. عدم وجود داده‌ها می‌تواند باعث انحراف در تجزیه و تحلیل آماری شود. علاوه بر این، بسیاری از الگوریتم‌های یادگیری ماشین از مجموعه داده‌هایی که دارای مقادیر گمشده هستند، پشتیبانی نمی‌کنند (Zhang & Thorburn, 2022). مرحله بعدی در آماده‌سازی مجموعه داده، شناسایی داده‌های گمشده و انتخاب رویکرد مناسب در برخورد با آنان است. بطور کلی برای مواجهه با داده‌های گمشده، دو رویکرد اصلی حذف یا جایگزینی، وجود دارد و به منظور جایگزینی داده‌های گمشده می‌توان از روش‌های ساده و یا مبتنی بر یادگیری ماشین استفاده کرد (Brownlee, 2020). در اغلب مواقع مقادیر داده‌ها برای پیاده‌سازی الگوریتم‌های یادگیری ماشین مناسب نیستند و بنابراین لازم است تبدیلاتی<sup>۳۰</sup> در ماهیت داده‌ها به منظور تطبیق مجموعه داده با نیازهای الگوریتم‌ها صورت بگیرد. مهمترین این تبدیلات، تبدیل متغیرهای اسمی و مقیاس‌بندی داده‌های مجموعه داده هستند (McCallum, 2012). در این پژوهش پس از جمع‌آوری داده‌ها و به منظور آماده‌سازی و مدل‌سازی، از زبان برنامه‌نویسی پایتون و محیط ژوپیتر نوت‌بوک<sup>۳۱</sup> استفاده شد. پس از فراخوانی مجموعه داده، داده‌های گمشده با میانگین<sup>۳۲</sup> جایگزین شدند. داده‌های کیفی با استفاده از روش تبدیل داده، کدگذاری دو دویی به ۰ و ۱ تبدیل شدند و سایر داده‌ها که ماهیت رتبه‌بندی داشتند، با استفاده از تبدیل رتبه‌بندی، کمی‌سازی شدند که این داده‌ها و تبدیلات انجام‌شده در پیوست ۱ قابل مشاهده است. پیش از پیاده‌سازی، لازم است تا داده‌ها به دو مجموعه داده آموزش و آزمون تقسیم شوند. این کار با استفاده از روش نگهدارنده<sup>۳۳</sup> با نسبت ۳۰ درصد انجام شد و مجموعه داده آزمون برای ارزیابی نهایی الگوریتم‌ها کنار گذاشته شد. سپس مجموعه داده حاصل با روش اعتبارسنجی متقابل<sup>۳۴</sup> (K=5) و با ۱۰۰ مرتبه تکرار، به دو مجموعه داده آموزش و اعتبارسنجی تقسیم شدند.

### ۳-۳) برخورد با عدم تعادل در داده‌های هدف مسئله

در بیشتر موارد، الگوریتم‌های رایج یادگیری ماشین در مسائل طبقه‌بندی، تعداد نمونه‌های کلاس‌های مختلف را مساوی فرض می‌کنند و در نتیجه، تعداد نابرابر داده‌ها در کلاس‌های مختلف باعث ایجاد مشکلاتی در فرآیند یادگیری الگوریتم‌ها

<sup>30</sup> Encoding

<sup>31</sup> Jupyter Notebook

<sup>32</sup> Median

<sup>33</sup> Holdout

<sup>34</sup> Repeated K-fold Cross Validation

می‌شود. از سوی دیگر در مسائل واقعی، جمع‌آوری مجموعه داده‌هایی که در آنها، داده‌ها در کلاس‌های مختلف به صورت متعادل توزیع شده باشند، دشوار است (Fotouhi et al., 2019). برای مواجهه با این مشکل، رویکردهای متفاوتی وجود دارد. یکی از این رویکردها، شبکه‌های متخاصم مولد<sup>۳۵</sup> هستند. این شبکه‌ها از دو شبکه مجزای مولد و تشخیص‌دهنده تشکیل می‌شوند. شبکه مولد، نویز تصادفی را به عنوان ورودی می‌گیرد و آن را به نمونه‌های مصنوعی تبدیل می‌کند. شبکه تشخیص‌دهنده مانند یک طبقه‌بندی‌کننده دودویی عمل کرده و بین نمونه‌های واقعی و مصنوعی تمایز قائل می‌شود. هر دو شبکه بطور همزمان در یک بازی مجموع صفر<sup>۳۶</sup> آموزش داده می‌شوند و در طول آموزش، شبکه مولد سعی می‌کند نمونه‌های مصنوعی تولید کند که به اندازه کافی واقعی باشد. این فرآیند تا زمانی ادامه می‌یابد که شبکه تشخیص‌دهنده فریب داده شود. فرآیند آموزش تا زمانی که شبکه مولد نمونه‌هایی را تولید کند که برای شبکه تشخیص‌دهنده از داده‌های واقعی قابل تمایز نباشند، ادامه می‌یابد (Goodfellow et al., 2020). علاوه بر این رویکرد، در این پژوهش برای غلبه بر عدم تعادل در داده‌های مسئله از روش‌های یادگیری گروهی و معیارهای ارزیابی متنوع (Ma & He, 2013) استفاده شده‌است.

### ۳-۴) مدل‌سازی

در مسائل یادگیری ماشین چهار رویکرد کلی یادگیری با نظارت<sup>۳۷</sup>، یادگیری بدون نظارت<sup>۳۸</sup>، یادگیری نیمه‌نظارتی<sup>۳۹</sup> و یادگیری تقویتی<sup>۴۰</sup> یا فعال<sup>۴۱</sup> وجود دارد. در پژوهش حاضر با توجه به اینکه داده‌ها دارای برچسب هستند و این برچسب‌ها گسسته هستند، از الگوریتم‌های طبقه‌بندی استفاده می‌شود (Han et al., 2012; Marsland, 2014). در این پژوهش ۹ الگوریتم طبقه‌بندی براساس چهار ویژگی پاسخگویی مناسب در برابر تعداد کم داده‌ها و عدم توازن در برچسب‌ها، مقاومت در برابر بیش‌برازش و پیچیدگی کم، مقایسه و با توجه به جدول (۲) چهار الگوریتم برتر، برای مدل‌سازی انتخاب شدند. در ارتباط با تعداد داده‌ها، پژوهش‌ها نشان می‌دهند که در مجموعه داده‌هایی با تعداد کم، الگوریتم‌های یادگیری ماشین نسبت به الگوریتم‌های یادگیری<sup>۴۲</sup> عمیق عملکرد بهتری از خود نشان می‌دهند (Bailly et al., 2022). از نظر توازن میان برچسب‌ها، طبقه‌بندی‌کننده‌های سنتی نظیر ماشین بردار پشتیبان و درخت تصمیم براساس توازن میان برچسب‌ها طراحی شده‌اند و در صورت عدم تعادل در برچسب‌های مسئله، تمایل به طبقه‌بندی نادرست و یا حذف طبقه اقلیت دارند (Gong & Kim, 2017). در مقابل، الگوریتم‌های یادگیری گروهی، عملکرد مناسبی در مجموعه داده‌هایی با عدم تعادل در برچسب‌ها دارند (Ding et al., 2023). در مورد ویژگی پیچیدگی و مستعد بیش‌برازش بودن الگوریتم‌ها لازم است از مفهوم تعمیم‌پذیری<sup>۴۳</sup> استفاده شود. تعمیم‌پذیری عبارت از توانایی مدل در به تصویر کشیدن ویژگی‌های جامعه با مشاهده ویژگی‌های نمونه است. زمانیکه آموزش تعمیم‌پذیر نباشد، مدل دچار بیش‌برازش خواهد شد (Myung, 2000). از آنجاییکه

<sup>35</sup> Generative adversarial networks (GANs)

<sup>36</sup> Zero-sum game

<sup>37</sup> Supervised Learning

<sup>38</sup> Unsupervised Learning

<sup>39</sup> Semi-supervised Learning

<sup>40</sup> Reinforcement Learning

<sup>41</sup> Active Learning

<sup>42</sup> Deep Learning

<sup>43</sup> Generalization

مجموعه داده مورد بررسی در این پژوهش، محدود و در نتیجه مستعد بیش‌برازش است، بهتر است برای جلوگیری از این اتفاق، تا حد امکان از الگوریتم‌هایی با پیچیدگی کمتر استفاده شود.

### جدول ۱۱. قابلیت الگوریتم‌های طبقه‌بندی

ویژگی‌های مسئله				الگوریتم‌های طبقه‌بندی
تعداد کم داده‌ها	عدم توازن در برچسب‌های مسئله	پیچیدگی کم	مقاومت در برابر بیش‌برازش	
*	*	*	*	
*	*	*	*	
*	*	*	*	
*	*	*	*	
*	*	*	*	
*	*	*	*	
*	*	*	*	
*	*	*	*	

اگرچه پژوهش‌ها با فرض‌های قانع‌کننده‌ای نظیر آنچه پیش از این گفته شد، به دنبال محدود کردن الگوریتم‌های قابل استفاده هستند اما نمی‌توان با قطعیت بر مبنای این دلایل، الگوریتمی را تا پیش از پیاده‌سازی، ارجح بر دیگری دانست. هرچند به‌علت محدودیت‌های موجود، در هر پژوهش معمولاً سعی می‌شود الگوریتم‌هایی را که ظاهراً عملکرد بهتری در مسئله مورد بررسی دارند، انتخاب و پیاده‌سازی شوند (Jafari, 2022). در ادامه چهار الگوریتم طبقه‌بندی که برای مدل‌سازی در این پژوهش استفاده شدند، شرح داده می‌شوند. اما لازم است پیش از آن، با توجه به ماهیت چندبرچسبی بودن مسئله، طبقه‌بندی این مسائل تبیین گردد. در مسائل چندبرچسبی، هدف مسئله شامل مجموعه‌ای از عوامل است. عمده‌ترین رویکردهای حل این مسائل، تبدیل مسئله<sup>۴۴</sup> و الگوریتم‌های تطابق‌یافته<sup>۴۵</sup> هستند (Ganda & Buch, 2018).

در تبدیل مسئله تلاش بر این است تا مسئله چندبرچسبی به مسئله‌های یک برچسبی تبدیل شود. رویکردهای حل مسئله با این روش عبارتند از: تطابق دودویی<sup>۴۶</sup>، زنجیره‌های طبقه‌بندی‌کننده<sup>۴۷</sup> و مجموعه توان برچسب‌ها<sup>۴۸</sup> که در تطابق دودویی، هر برچسب به‌عنوان یک مسئله مستقل تک‌برچسبی در نظر گرفته می‌شود. در روش زنجیره‌های طبقه‌بندی‌کننده، منحصر با یک برچسب آموزش می‌بیند و طبقه‌بند بعدی روی طبقه‌بند قبلی و برچسب بعدی آموزش می‌بینند. در مجموعه توان برچسب‌ها، تلاش بر این است تا مسئله چندبرچسبی به یک مسئله چندکلاسه<sup>۴۹</sup> تبدیل شود (Pushpa & Karpagavalli, 2017). در الگوریتم‌های تطابق‌یافته، از الگوریتم‌هایی برای حل مسئله استفاده می‌شود که صرفاً برای حل

<sup>44</sup> Problem Transformation

<sup>45</sup> Adapted Algorithm

<sup>46</sup> Binary Relevance

<sup>47</sup> Classifier Chains

<sup>48</sup> Label Powerset

<sup>49</sup> Multi-Class

اینگونه مسائل توسعه یافته‌اند (Zheng et al., 2019). در ادامه چهار الگوریتم تطابق‌یافته‌ای که برای مدل‌سازی در این پایان‌نامه استفاده شدند، بطور مختصر شرح داده می‌شود.

### ۳-۴-۱) K- نزدیکترین همسایه<sup>۵۰</sup>

در این الگوریتم از یک سنجه<sup>۵۱</sup> فاصله مانند فاصله اقلیدسی<sup>۵۲</sup> برای یافتن K نزدیک‌ترین نقاط به نقطه هدف استفاده می‌شود و پس از آن برجسب را براساس بیشترین برجسب در میان نقاط، به نقطه هدف اختصاص می‌دهد (Han et al., 2012).

### ۳-۴-۲) تقویت گرادیان شدید<sup>۵۳</sup>

این الگوریتم یکی از الگوریتم‌های توسعه یافته یادگیری گروهی<sup>۵۴</sup> با رویکرد تقویت کردن<sup>۵۵</sup> است که بر مبنای گرادیان کاهش<sup>۵۶</sup> عمل می‌کند (Dong et al., 2020). این الگوریتم با یک مدل ساده مبتنی بر درخت تصمیم شروع می‌شود و سپس درخت‌های دیگر را با رویکرد حریصانه و با استفاده از تقویت گرادیان، تکنیک‌های منظم‌سازی و هرس بهینه می‌کند. در این روش در هر مرحله از تکرار میزان پیچیدگی درخت به صورت جداگانه انتخاب و کنترل می‌شود و تعداد برگ‌ها در همه تکرارها ثابت نخواهد بود (Chen et al., 2015).

### ۳-۴-۳) تقویت دسته<sup>۵۷</sup>

تقویت دسته (Dorogush et al., 2018) از زیرمجموعه رویکردهای تقویت گرادیان الگوریتم‌های یادگیری گروهی است. این الگوریتم شامل تکنیک‌های منظم‌سازی برای جلوگیری از بیش‌برازش است و همچنین قابلیت آموزش موازی را برای محاسبات سریع‌تر فراهم می‌کند. این الگوریتم از دو مرحله برای ساخت درخت تشکیل شده است. مرحله اول عبارتست از انتخاب ساختار درخت و مرحله دوم شامل تعیین مقدار برگ است (Tanha et al., 2020).

### ۳-۴-۴) جنگل تصادفی<sup>۵۸</sup>

در این الگوریتم که یکی از الگوریتم‌های یادگیری گروهی است، چندین درخت تصمیم بر روی زیرمجموعه‌های مختلف داده‌های آموزشی و زیرمجموعه‌های مختلف ویژگی‌های ورودی، آموزش داده می‌شوند. هر درخت باید روی همان زیرمجموعه، طبقه‌بند خود را بسازد و پیش‌بینی را ارائه دهد. پیش‌بینی نهایی تجمیع اکثریت آرای پیش‌بینی‌های درختان تصمیم‌گیر منفرد است. جنگل‌های تصادفی با وجود سادگی، یکی از قدرتمندترین الگوریتم‌های یادگیری ماشین در دسترس هستند و به دلیل ترکیب نتایج درختان مختلف احتمال پایداری وجود دارد تا دچار بیش‌برازش نشوند (Polikarov et al., 2019).

<sup>50</sup> K-Nearest Neighbors (KNN)

<sup>51</sup> Metric

<sup>52</sup> Euclidean Distance

<sup>53</sup> XGBoost (Extreme Gradient Boosting)

<sup>54</sup> Ensemble Learning

<sup>55</sup> Boosting

<sup>56</sup> Gradient descent

<sup>57</sup> Category Boosting (CatBoost)

<sup>58</sup> Random Forrest

### ۳-۵) ارزیابی

پیش از ارزیابی عملکرد الگوریتم‌های پیاده‌سازی شده، لازم است مجموعه داده به دو مجموعه داده آموزش و ارزیابی تقسیم شوند. در این پژوهش ابتدا از روش نگهدارنده<sup>۵۹</sup> برای کنار گذاشتن بخشی از مجموعه داده به عنوان مجموعه داده آزمون استفاده و سپس روش اعتبارسنجی<sup>۶۰</sup> متقابل برای تقسیم مجموعه داده آموزش به دو مجموعه آموزش و اعتبارسنجی بکار برده می‌شود.

### ۳-۵-۱) معیارهای ارزیابی عملکرد الگوریتم‌های ارزیابی

معیارهای ارزیابی مورد استفاده در این پژوهش و نحوه محاسبه آنها در جدول (۳) ارائه شده‌اند. معیارهای دقت و یادآوری از ماتریس درهم‌ریختگی<sup>۶۱</sup> استخراج می‌شوند (Han et al., 2012) و معیار شاخص ارزیابی F1، میانگین هارمونیک دو معیار بازخوانی و دقت است (Rivolli et al., 2020).

جدول ۱۲. معیارهای ارزیابی

نام معیار	رابطه
دقت <sup>۶۲</sup>	$\text{دقت} = \frac{\text{تعداد منفی صحیح} + \text{تعداد مثبت صحیح}}{\text{تعداد منفی کاذب} + \text{تعداد مثبت صحیح}}$
بازخوانی <sup>۶۳</sup>	$\text{بازخوانی} = \frac{\text{تعداد مثبت صحیح}}{\text{تعداد منفی صحیح} + \text{تعداد مثبت صحیح}}$
شاخص ارزیابی F1 <sup>۶۴</sup> (Tsoumakas et al., 2010)	$F1 - \text{Score} = 2 \times \frac{\text{بازخوانی} \times \text{دقت}}{\text{بازخوانی} + \text{دقت}}$
ضرر همینگ <sup>۶۵</sup> (Wu & Zhu, 2020)	$\text{Hamming Loss} = \frac{1}{NL} \sum_{j=1}^N \sum_{i=1}^L \text{xor}[\tilde{y}_j^i \neq y_j^i]$
	<p><math>y_j^i</math> = برچسب‌های واقعی</p> <p><math>\tilde{y}_j^i</math> = برچسب‌های پیش‌بینی شده</p> <p><math>N</math> = تعداد نمونه‌ها یا رکوردها</p>
جاکارد <sup>۶۶</sup> (Krawezik et al., 2018)	$\text{Jaccard Index} = \frac{\tilde{y}_j^i \cap y_j^i}{\tilde{y}_j^i \cup y_j^i}$

<sup>59</sup> Hold-Out

<sup>60</sup> Cross Validation

<sup>61</sup> Confusion matrix

<sup>62</sup> Precision

<sup>63</sup> Recall

<sup>64</sup> F1-Score

<sup>65</sup> Hamming Loss

<sup>66</sup> Jaccard

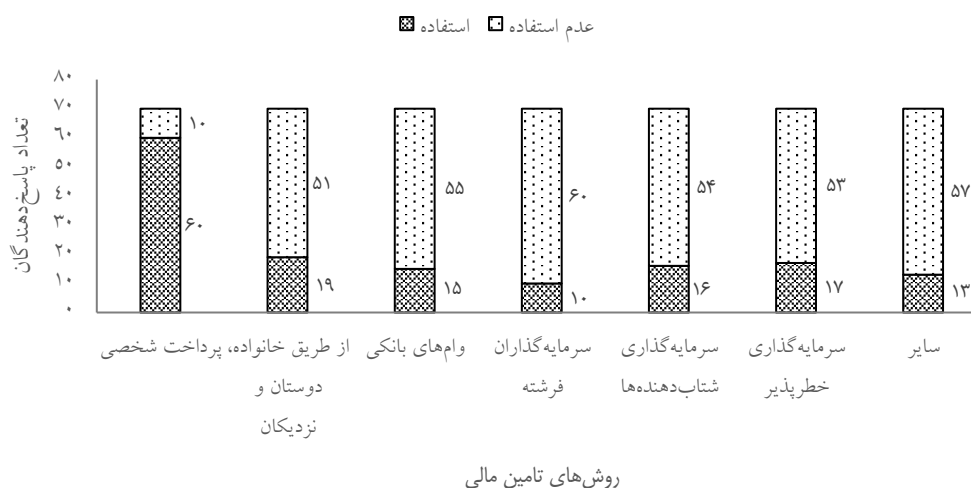
#### ۴ یافته‌های پژوهش

پیش از آنکه ادامه مراحل مدل‌سازی انتخاب شیوه تامین مالی در استارت‌آپ‌ها با در نظر گرفتن الگوهای رفتاری تشریح گردد، لازم است تا شناختی از مجموعه داده مورد بررسی به دست آید. به این منظور در بخش بعدی، داده‌های حاصل از پرسشنامه به اختصار تشریح می‌گردند.

#### ۴-۱ آمار توصیفی

همانطور که گفته شد تعداد داده‌های جمع‌آوری شده در این پژوهش ۷۰ نفر از بنیان‌گذاران استارت‌آپ‌ها، شامل ۵۶ مرد و ۱۴ زن بودند که ۷۴ درصد آنان در بازه سنی کمتر از ۴۰ سال قرار دارند. این مسئله حاکی از این امر است که بیشتر بنیان‌گذاران استارت‌آپ‌ها جوان هستند و در تصمیم‌گیری‌ها از راهنمایی‌های بیرونی استفاده می‌نمایند. ۳۲ درصد بنیان‌گذاران پاسخ‌دهنده به پرسشنامه این پژوهش، درآمدی کمتر از ۱۰ میلیون تومان در ماه دارند. توجه به این مسئله برای انتخاب روش‌های تامین مالی نکته بااهمیتی است چراکه درآمد بالا تنها عاملی است که افراد را در مراحل ابتدایی و راه‌اندازی، از اتکا به روش‌های تامین مالی بیرونی بی‌نیاز کرده و همچنین در بسیاری از روش‌های تامین مالی، یکی از ملاک‌های تاثیرگذار در اعتبارسنجی مالی است. حدود ۸۳ درصد پاسخ‌دهندگان دارای تحصیلات کارشناسی یا کارشناسی‌ارشد هستند که در دسترسی ایشان به شیوه‌های تامین مالی مختلف موثر است.

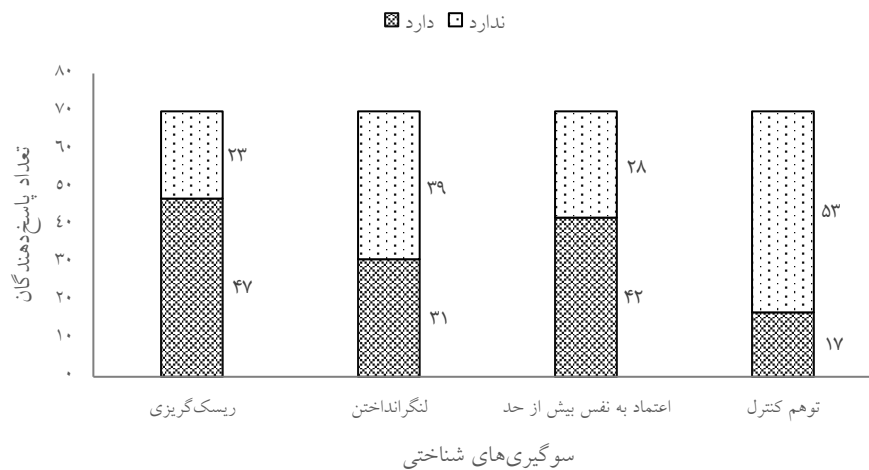
مطابق با پژوهش‌های پیشین، در دسترس‌ترین شیوه تامین مالی برای استارت‌آپ‌ها در مراحل اولیه، روش پرداخت شخصی پیش‌بینی شد و نتایج داده‌های این پژوهش نیز موید و حاکی از این است که نزدیک به ۸۶ درصد بنیان‌گذاران استارت‌آپ‌ها از این روش جهت تامین مالی استفاده کرده‌اند.



شکل ۲۷. روش‌های تامین مالی که بنیان‌گذاران استارت‌آپ‌ها استفاده/عدم استفاده کرده‌اند

نتایج حاصل از چهار سوگیری رفتاری که در پرسشنامه این پژوهش مورد سنجش قرار گرفتند، در شکل (۳) قابل مشاهده است. به ترتیب بیشترین سوگیری‌های مشاهده شده در بین بنیان‌گذاران پاسخ‌دهنده به پرسشنامه این پژوهش، ریسک‌گریزی،

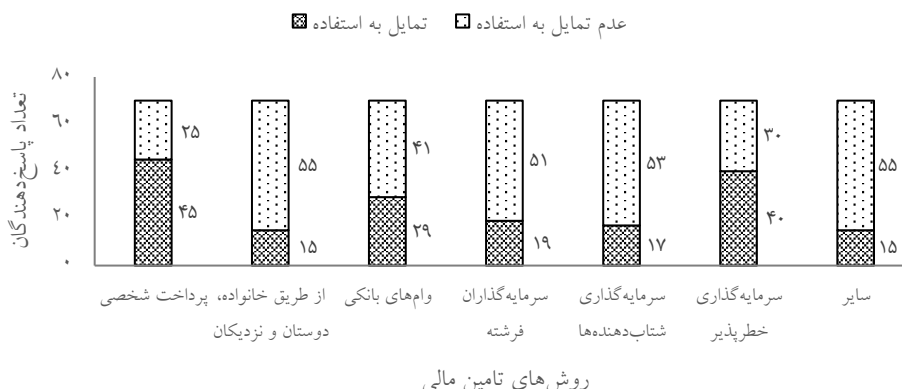
اعتماد به نفس بیش از حد، لنگرانداختن و توهم کنترل است. این مسئله تایید کننده این است که بنیان گذاران استارت‌آپ‌ها در معرض سوگیری‌های رفتاری هستند. برخلاف تصور اقتصاد متعارف از رفتار انسان در تصمیمات مالی و غیرمالی، سوگیری‌های رفتاری موجب فاصله گرفتن افراد از تصمیمات عقلانی می‌شود. اطلاع از این امر که افراد در تصمیم‌گیری عقلایی عمل نمی‌کنند، برای سرمایه‌گذاران و سرمایه‌پذیران این فرصت را فراهم می‌کند که با به دست آوردن شناخت کافی از این سوگیری‌ها از آثار منفی آن در تصمیمات خود اجتناب کنند.



شکل ۳. سوگیری‌های شناسایی شده در بنیان‌گذاران پاسخ‌دهنده به پرسشنامه پژوهش

شکل (۴) نشان‌دهنده روش‌های تامین مالی است که بنیان‌گذاران استارت‌آپ‌ها در صورت فراهم بودن تمامی شرایط، تمایل به استفاده از آن دارند. مطابق این نمودار حتی در صورت فراهم بودن تمامی شرایط، روش پرداخت شخصی اولویت اصلی پاسخ‌دهندگان را تشکیل می‌دهد. این مسئله تایید کننده و همسو با فراوانی سوگیری زیان‌گریزی در بین پاسخ‌دهندگان این پرسشنامه است و نشان می‌دهد که لزوماً در دسترس بودن روش تامین مالی، علت انتخاب یا عدم انتخاب یک روش نیست. نکته قابل توجه دیگر در این نمودار، گزینه سایر در بین روش‌های تامین مالی است. مطابق پاسخ‌ها، سایر روش‌های تامین مالی شامل کمک‌های حمایتی نهادهایی مانند صندوق نوآوری و شکوفایی، معاونت علمی، فناوری و اقتصاد دانش‌بنیان ریاست جمهوری، پارک‌های علم و فناوری، مراکز رشد و صندوق‌های پژوهش و فناوری است. این روش‌های تامین مالی معمولاً دارای کارمزدهای ترجیحی نسبت به سایر روش‌های تامین مالی هستند و شرایط دریافت آنان به گونه‌ای است که برای استارت‌آپ‌ها قابل استفاده باشند.





شکل ۴. روش‌های تامین مالی که بنیان‌گذاران پاسخ‌دهنده در این پژوهش تمایل به استفاده/عدم استفاده داشتند

#### ۴-۲) مدل‌سازی و پیاده‌سازی الگوریتم‌ها

همانطور که در بخش قبل شرح داده شد، الگوریتم‌ها بر روی مجموعه داده پیاده‌سازی شدند و نتایج پیاده‌سازی الگوریتم‌ها بر روی مجموعه داده اعتبارسنجی در جدول (۴) قابل مشاهده است.

جدول ۱۳. عملکرد الگوریتم‌های پیاده‌سازی بر روی مجموعه داده اعتبارسنجی

جاکارد	ضرر همینگ	شاخص ارزیابی F1	فراخوانی	دقت	رویکرد حل مسئله	
۰/۳۷	۰/۵۵	%۵۲	%۴۰	%۸۲	تطابق دو دویی	تبدیل مسئله
۰/۴۶	۰/۴۷	%۶۱	%۵۰	%۸۴	زنجیره‌های طبقه‌بندی کننده	
۰/۷۵	۰/۲۳	%۸۵	%۸۵	%۸۵	مجموعه توان برچسب‌ها	
۰/۷۸	۰/۲۱	%۸۸	%۹۳	%۸۵	تقویت گرادیان شدید	الگوریتم‌های تطابق یافته
۰/۸	۰/۱۹	%۸۹	%۹۵	%۸۴	K-نزدیکترین همسایه (K=۱۳)	
۰/۸۳	۰/۱۵	%۹۱	%۹۹	%۸۴	تقویت دسته	
۰/۸۱	۰/۱۸	%۹۰	%۹۸	%۸۴	جنگل تصادفی	

مشاهده می‌شود که اکثر الگوریتم‌های پیاده‌سازی شده، عملکرد قابل قبولی را با معیارهای ارزیابی در مجموعه داده اعتبارسنجی ارائه داده‌اند. این امر نشان می‌دهد که انتخاب عوامل و رویکرد حل مسئله با توجه به ماهیت داده‌ها به درستی طی شده است. در میان الگوریتم‌های پیاده‌سازی شده، الگوریتم تقویت دسته در اکثریت معیارهای ارزیابی، عملکرد بهتری نسبت به سایر الگوریتم‌ها در مجموعه داده اعتبارسنجی دارد. بنابراین الگوریتم تقویت دسته انتخاب شده و مجموعه داده آزمون که کنار گذاشته شده بود، فراخوانی می‌گردد. پس از آن آماده‌سازی داده‌ها اجرا شده و الگوریتم بر روی آن پیاده‌سازی می‌گردد. در جدول (۵) عملکرد این الگوریتم در مجموعه داده آزمون قابل مشاهده است.

جدول ۱۴. ارزیابی عملکرد بهترین الگوریتم پیاده‌سازی در مجموعه داده آزمون

جاکارد	ضرر همینگ	شاخص ارزیابی F1	فراخوانی	دقت	رویکرد حل مسئله
۰/۷۹	۰/۲۰	%۸۹	%۹۹	%۸۱	تقویت دسته

مقادیر ویژگی‌های ورودی و ویژگی‌های پیش‌بینی‌شده برای مجموعه داده آزمون این پژوهش به شرح پیوست ۲ است. همانطور که در جدول (۵) مشاهده می‌شود با استفاده از این الگوریتم، مدل قادر خواهد بود با دریافت ورودی‌ها که مطابق توضیحات بخش ۳ آماده‌سازی می‌شوند، روش‌های تامین مالی که بنیان‌گذاران استارت‌آپ‌ها انتخاب می‌کنند را در مجموعه داده آزمونی که در فرآیند آموزش حضور نداشته‌است، با شاخص ارزیابی F1 معادل ۸۹ درصد پیش‌بینی نماید.

#### ۴-۲-۱) مطالعه موردی

برای بررسی عملکرد مدل؛ یک بنیان‌گذار استارت‌آپ مراجعه‌کننده برای تامین مالی، با ویژگی‌های ورودی مطابق جدول (۶) مورد مطالعه قرار گرفته‌است. پس از طی فرآیند آماده‌سازی داده‌ها، نتیجه پیش‌بینی در جدول (۷) نشان داده شده‌است. مطابق نتیجه به‌دست آمده، فرد مورد مطالعه از روش‌های تامین مالی پرداخت شخصی و سرمایه‌گذاری خطرپذیر استفاده می‌کند. نتیجه حاصل از مطالعه استارت‌آپ‌های مختلف به سرمایه‌گذاران اعم از شتاب‌دهنده‌ها، پارک‌های علم و فناوری، سکوها تامین مالی و سایرین امکان می‌دهد که با در اختیار داشتن این اطلاعات بتوانند در مذاکرات تامین مالی، پیشنهادات مناسبی به سرمایه‌پذیران ارائه دهند.

#### جدول ۶. نمونه داده ورودی

سویچ توییتم	سویچ اعتماد به نفس	سویچ لنگر انداختن	سویچ زبان‌گزینی	سرمایه‌گذاران خطرپذیر شرکتی	سرمایه‌گذاران خطرپذیر	شتاب‌دهنده‌ها	سرمایه‌گذاران وام‌های بانکی	خانواده و دوستان	توهم	سرمایه کنترل
۱	۰	۱	۱	۱	۱	۰	۱	۰	۰	۱

سابقه فروش	سابقه تولید	ثابت اختراع	نوع استارت‌آپ	تعداد اعضای تیم	سابقه کار کارآفرینی	سابقه کار تجاری	درآمد ماهانه	تحصیلات	سن	جنسیت	سن استارت‌آپ	فعال استارت‌آپ؟
۱	۱	۱	۰	۱۲	۱۲	۰	۵	۵	۴	۱	۵	۱

#### جدول ۷. روش‌های تامین مالی پیش‌بینی‌شده توسط مدل توسعه یافته برای نمونه داده ورودی

سرمایه‌گذاران خطرپذیر شرکتی	سرمایه‌گذاران خطرپذیر	سرمایه‌گذاران وام‌های بانکی	خانواده و دوستان	توهم شخصی
۱	۰	۱	۱	۰

#### ۵) نتیجه‌گیری و پیشنهادها

این پژوهش با هدف پیش‌بینی انتخاب شیوه تامین مالی در استارت‌آپ‌ها، با ادغام سوگیری‌های رفتاری در چارچوب مدل‌سازی داده‌محور، بازنمایی جامع‌تر و واقعی‌تر از فرآیند تصمیم‌گیری را ارائه می‌دهد. همچنین کارآفرینان و سرمایه‌گذاران را قادر می‌سازد تا پیچیدگی‌های تامین مالی استارت‌آپ‌ها را به‌طور موثرتری دنبال کنند. در این پژوهش ۱۰ روش تامین مالی رایج در ادبیات موضوع شناسایی شده و تعدادی از عوامل موثر و در دسترس، برای مدل‌سازی انتخاب عملکرد مدل بر روی مجموعه داده آموزش و آزمون مورد بررسی قرار گرفت. همانطور که مشاهده شد مدل توسعه یافته با

دقت قابل قبولی، روش‌های تامین مالی مناسب برای هر استارت‌آپ را پیش‌بینی می‌کند و به سرمایه‌گذاران امکان می‌دهد تا پیشنهادات شخصی‌سازی شده‌ای ارائه دهند. با بررسی داده‌های جمع‌آوری شده مشاهده گردید که فراوانی سوگیری‌های رفتاری در بنیان‌گذاران استارت‌آپ‌ها با انتخاب شیوه‌های تامین مالی که استفاده یا تمایل به استفاده دارند، هم‌خوانی دارد. نتایج این پیش‌بینی می‌تواند برای سیاست‌گذاران جهت تدوین برنامه و فراهم آوردن زیرساخت‌های مناسب بابت استفاده از شیوه‌های تامین مالی مختلف و متناسب با نیاز هریک از کارآفرینان موثر باشد. همچنین این نتایج برای سرمایه‌گذاران جهت شناسایی انتخاب‌ها و اولویت‌های کارآفرینان و بهره‌برداری از آن بابت دستیابی به اهداف سرمایه‌گذاری کمک‌کننده می‌باشد. این موضوع برای کارآفرینان نیز بی‌تاثیر نبوده و جهت سنجش گزینه‌های تامین مالی پیش از ورود به بازار و راه‌اندازی کسب‌وکار و یا در جهت توسعه کسب‌وکار مفید خواهد بود.

در میان روش‌های استفاده‌شده برای مدل‌سازی در این پژوهش، الگوریتم تقویت دسته از الگوریتم‌های یادگیری گروهی می‌تواند روش‌های تامین مالی که استارت‌آپ‌ها انتخاب می‌کنند را با شاخص ارزیابی F1 معادل ۸۹ درصد و دقت معادل ۸۵ درصد بر روی مجموعه داده آزمون پیش‌بینی نماید. برای ادامه این پژوهش می‌توان مدل‌سازی را با در نظر گرفتن عوامل اقتصاد کلان از جمله انتظارات تورمی و بررسی تمایل به ریسک‌پذیری افراد با تغییرات انتظارات تورمی انجام داد و یا مدل این مسئله را در موسسات مالی پیاده‌سازی کرده و صحت عملکرد آن را بررسی نمود. همچنین با توجه به استفاده بالای ۸۵ درصدی از روش تامین مالی پرداخت شخصی در میان بنیان‌گذاران استارت‌آپ‌های این پژوهش، در پژوهش‌های آتی به واکاوی این انتخاب پرداخته می‌شود.

## منابع

- Azoulay, P., Jones, B. F., Kim, J. D., & Miranda, J. (2020). Age and high-growth entrepreneurship. *American Economic Review: Insights*, 2(1), 65-82. <https://dx.doi.org/10.3386/w24489>
- Bailly, A., Blanc, C., Francis, É., Guillotin, T., Jamal, F., Wakim, B., & Roy, P. (2022). Effects of dataset size and interactions on the prediction performance of logistic regression and deep learning models. *Computer Methods and Programs in Biomedicine*, 213, 106504. <http://dx.doi.org/10.1016/j.cmpb.2021.106504>.
- Bazerman, M. H., & Moore, D. A. (2012). Judgment in managerial decision making. John Wiley & Sons. <http://dx.doi.org/10.4324/9780203141939-11>
- Bolarinwa, O. A. (2015). Principles and methods of validity and reliability testing of questionnaires used in social and health science researches. *Nigerian Postgraduate Medical Journal*, 22(4), 195. <http://dx.doi.org/10.4103/1117-1936.173959>
- Brownlee, J. (2020). Data preparation for machine learning: data cleaning, feature selection, and data transforms in Python. *Machine Learning Mastery*.
- Cassar, G. (2004). The financing of business start-ups. *Journal of Business Venturing*, 19(2), 261-283. [https://doi.org/10.1016/S0883-9026\(03\)00029-6](https://doi.org/10.1016/S0883-9026(03)00029-6)
- Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., Chen, K., Mitchell, R., Cano, I., & Zhou, T. (2015). Xgboost: extreme gradient boosting. R package version 0.4-2, 1(4), 1-4. <http://dx.doi.org/10.32614/cran.package.xgboost>
- Ding, H., Sun, Y., Wang, Z., Huang, N., Shen, Z., & Cui, X. (2023). RGAN-EL: A GAN and ensemble learning-based hybrid approach for imbalanced data classification. *Information Processing & Management*, 60(2), 103235. <http://dx.doi.org/10.1016/j.ipm.2022.103235>
- Dominic, C., & Gupta, A. (2020). Psychological factors affecting investors decision making. *Journal of Xi'an University of Architecture and Technology*, 7(6), 169-181. <http://dx.doi.org/10.55041/ijsrem30872>.
- Dong, X., Yu, Z., Cao, W., Shi, Y., & Ma, Q. (2020). A survey on ensemble learning. *Frontiers of Computer Science*, 14, 241-258. <http://dx.doi.org/10.1007/s11704-019-8208>.
- Dorogush, A. V., Ershov, V., & Gulin, A. (2018). CatBoost: gradient boosting with categorical features support. arXiv preprint arXiv:1810.11363. <https://doi.org/10.48550/arXiv.1810.11363>.

- Elston, D. M. (2021). Survivorship bias. *Journal of the American Academy of Dermatology*. <http://dx.doi.org/10.1016/j.jaad.2021.06.84> .
- Fotouhi, S., Asadi, S., & Kattan, M. W. (2019). A comprehensive data level analysis for cancer diagnosis on imbalanced data. *Journal of biomedical informatics*, 90, 103089. <https://doi.org/10.1016/j.jbi.2018.12.003>
- Franco, S., Cappa, F., & Pinelli, M. (2021). Founder Education and Start-Up Funds Raised. *IEEE Engineering Management Review*, 49(3), 42-48. <https://doi.org/10.1109/EMR.2021.3077966>
- Frid, C. J., Wyman, D. M., Gartner, W. B., & Hechavarria, D. H. (2016). Low-wealth entrepreneurs and access to external financing. *International Journal of Entrepreneurial Behavior & Research*. <http://dx.doi.org/10.1108/ijeb-08-2015-0173>
- Ganda, D., & Buch, R. (2018). A survey on multi label classification. *Recent Trends in Programming Languages*, 5(1), 19-23
- Gong, J., & Kim, H. (2017). RHSBoost: Improving classification performance in imbalance data. *Computational Statistics & Data Analysis*, 111, 1-13. <http://dx.doi.org/10.1016/j.csda.2017.01.005>
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2020). Generative adversarial networks. *Communications of the ACM*, 63(11), 139-144. <https://doi.org/10.1145/3422622>
- Han, J., Kamber, M., & Pei, J. (2012). *Data mining concepts and techniques third edition*. University of Illinois at Urbana-Champaign Micheline Kamber Jian Pei Simon Fraser University .
- Hebert, C. (2020). Gender stereotypes and entrepreneur financing. 10th Miami Behavioral Finance Conference.
- Israel, G. D. (1992). Determining sample size. <https://dx.doi.org/10.2139/ssrn.3318245>
- Jafari, R. (2022). *Hands-On Data Preprocessing in Python: Learn How to Effectively Prepare Data for Successful Data Analytics*. Packt Publishing. <https://books.google.com/books?id=nzmznzgEACAAJ>.
- Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan .
- Krawezik, G. P., Kogge, P. M., Dysart, T. J., Kuntz, S. K., & McMahon, J. O. (2018). Implementing the jaccard index on the migratory memory-side processing emu architecture. 2018 IEEE High Performance extreme Computing Conference (HPEC).
- Langer, E. J. (1975). The illusion of control. *Journal of personality and social psychology*, 32(2), 311. <http://dx.doi.org/10.1037//0022-3514.32.2.311>
- Lybaert, N., & Umans, I. (2022). Start-up Performance: Looking for an Explanation in Entrepreneurial Characteristics and Financing Choice. *European Conference on Innovation and Entrepreneurship*, <https://doi.org/10.34190/ecie.17.1.833>
- Marsland, S. (2014). *Machine Learning: An Algorithmic Perspective, Second Edition*. CRC Press. <https://books.google.com/books?id=6GvSBQAAQBAJ>.
- McCallum, Q. E. (2012). *Bad data handbook: cleaning up the data so you can get back to work*. " O'Reilly Media, Inc .".
- Myung, I. J. (2000). The importance of complexity in model selection. *Journal of mathematical psychology*, 44(1), 190-204. <https://doi.org/10.1006/jmps.1999.1283>
- Ouimet, P., & Zarutskie, R. (2014). Who works for startups? The relation between firm age, employee age, and growth. *Journal of financial Economics*, 112(3), 386-407. <https://doi.org/10.1016/j.jfineco.2014.03.003>
- Pavlov, Y. L. (2019). Random forests. In *Random Forests*. De Gruyter .
- Pushpa, M., & Karpagavalli, S. (2017). Multi-label classification: problem transformation methods in Tamil phoneme classification. *Procedia Computer Science*, 115, 572-579. <https://doi.org/10.1016/j.procs.2017.09.116>
- Ramalakshmi, V., Pathak, V. K., & Mary, C. (2019). Impact of Cognitive Biases on investment decision making. *Journal of Critical Reviews*, 6(6), 59-64
- Rivoli, A., Read, J., Soares, C., Pfahringer, B., & de Carvalho, A. C. (2020). An empirical analysis of binary transformation strategies and base algorithms for multi-label learning. *Machine Learning*, 109, 1509-1563. <https://doi.org/10.1007/s10994-020-05879-3>
- Rosenfeld, A., & Kraus, S. (2018). Predicting human decision-making: From prediction to action. *Synthesis lectures on artificial intelligence and machine learning*, 12(1), 1-150. [https://doi.org/10.1007/978-3-031-01578-6\\_3](https://doi.org/10.1007/978-3-031-01578-6_3)
- Rosyidah, U., & Pratikto, H. (2022). The role of behavioral bias on financial decision making: a systematic literature review and future research agenda. *Journal of Enterprise and Development (JED)*, 4(1), 156-179. <https://doi.org/10.20414/jed.v4i1.5102>
- Simon, M., & Houghton, S. M. (2003). The relationship between overconfidence and the introduction of risky products: Evidence from a field study. *Academy of management journal*, 46(2), 139-149. <https://doi.org/10.5465/30040610>
- Tanha, J., Abdi, Y., Samadi, N., Razzaghi, N., & Asadpour, M. (2020). Boosting methods for multi-class imbalanced data classification: an experimental review. *Journal of Big Data*, 7, 1-47. <https://doi.org/10.1186/s40537-020-00349-y>
- Tech, R. P. (2018). *Financing high-tech startups*. Springer. <https://doi.org/10.1007/978-3-319-66155-1>
- Tsoumakas, G., Katakis, I., & Vlahavas, I. (2010). Mining multi-label data. *Data mining and knowledge discovery handbook*, 667-685. <https://doi.org/10.1007/978-0-387-09823-434>

- Ul Abidin, S. Z., Qureshi, F., Iqbal, J., & Sultana, S. (2022). Overconfidence bias and investment performance: A mediating effect of risk propensity. *Borsa Istanbul Review*, 22(4), 780-793. <https://doi.org/10.1016/j.bir.2022.03.001>
- Vo, D. H. (2019). Patents and Early-Stage Financing: Matching versus Signaling. *Journal of small business management*, 57(4), 1252-1279. <https://doi.org/10.1111/jsbm.12414>
- Wu, G., & Zhu, J. (2020). Multi-label classification: do Hamming loss and subset accuracy really conflict with each other? *Advances in Neural Information Processing Systems*, 33, 3130-3140. <https://doi.org/10.48550/arXiv.2011.07805>
- Zahera, S. A., & Bansal, R. (2018). Do investors exhibit behavioral biases in investment decision making? A systematic review. *Qualitative Research in Financial Markets*. <https://doi.org/10.1108/QRFM-04-2017-0028>
- Zhang, S. X., & Cueto, J. (2017). The study of bias in entrepreneurship. *Entrepreneurship theory and Practice*, 41(3), 419-454. <https://doi.org/10.1111/etap.12212>
- Zhang, S. X., Foo, M.-D., & Vassolo, R. S. (2021). The ramifications of effectuation on biases in entrepreneurship—Evidence from a mixed-method approach. *Journal of Business Venturing Insights*, 15, e00238. <https://doi.org/10.1016/j.jbvi.2021.e00238>
- Zhang, Y., & Thorburn, P. J. (2022). Handling missing data in near real-time environmental monitoring: A system and a review of selected methods. *Future Generation Computer Systems*, 128, 63-72. <https://doi.org/10.1016/j.future.2021.09.033>

## پیوست (۱)

## جدول ۷. متغیرهای مستقل و هدف پژوهش

ردیف	نام ویژگی	مقدار
۱	آیا شرکت شما فعال است؟	بله
		خیر
۲	چند سال از تاسیس شرکت شما می‌گذرد؟	
۳	جنسیت	زن
		مرد
۴	سن	زیر ۲۵
		۲۶ تا ۳۰
		۳۱ تا ۳۵
		۳۶ تا ۴۰
۵	تحصیلات	زیر دیپلم
		دیپلم
		فوق دیپلم
		کارشناسی
		کارشناسی ارشد
		دکتری
۶	میزان درآمد ماهانه	درآمدی ندارد
		زیر ۵م
		از ۵ تا ۱۰م
		از ۱۰ تا ۱۵م
		از ۱۵ تا ۲۰م
		۲۰م به بالا
		تمایلی ندارد بگویم

		سابقه کار تجاری	۷
		سابقه کار آفرینی	۸
		تعداد کارمندان شرکت	۹
۰	خدماتی	نوع شرکت	۱۰
۱	تولیدی		
۲	سایر		
۰	پایین	درجه فناوری	۱۱
۱	متوسط		
۲	بالا		
۱	دارد	ثبت اختراع	۱۲
۰	ندارد		
۱	دارد	تولید محصول	۱۳
۰	ندارد		
۱	دارد	فروش	۱۴
۰	ندارد		
		میزان سرمایه راه اندازی	۱۵
در صورت امکان تمایل به استفاده از چه شیوه/های تامین مالی دارید؟ ۱: تمایل ندارد و ۰: تمایل دارد		سرمایه شخصی	۱۶
		خانواده و دوستان	۱۷
		وام های بانکی	۱۸
		سرمایه گذاران فرشته	۱۹
		شتاب دهنده ها	۲۰
		سرمایه گذاران خطر پذیر	۲۱
		سرمایه گذاران خطر پذیر شرکتی	۲۲
		سایر	۲۳
۱	دارد	سوگیری اعتماد به نفس بیش از حد	۲۴
۰	ندارد		
۱	دارد	سوگیری زیان گریزی	۲۵
۰	ندارد		
۱	دارد	سوگیری لنگر انداختن	۲۶
۰	ندارد		
۱	دارد	سوگیری توهم کنترل	۲۷
۰	ندارد		
استفاده از شیوه/های تامین مالی؟ ۱: استفاده نمی کند و ۰: استفاده می کند		پرداخت شخصی	۲۸
		خانواده و دوستان	۲۹
		وام های بانکی	۳۰
		سرمایه گذاران فرشته	۳۱
		شتاب دهنده ها	۳۲

سرمایه‌گذاران خطرپذیر	۳۳
سرمایه‌گذاران خطرپذیر شرکتي	۳۴
سایر	۳۵

**پیوست (۲)**

مقادیر ویژگی‌های مستقل و هدف در جدول ۹ و ۱۰ قابل مشاهده هستند.

**جدول ۹. مجموعه داده آزمون (ویژگی‌های ورودی)**

نمونه‌ها	استارت‌آپ فعال	سن استارت‌آپ	جنسیت	سن	تجصیلات	درآمد ماهانه	سابقه کار تجاری	سابقه کار	تعداد اعضای تیم	نوع استارت‌آپ	ثبت اختراع	سابقه تولید	سابقه فروش	سرمایه اولیه	پرداخت شخصی	خانواده و	وام‌های بانکی	سرمایه‌گذاران	شباب‌دهنده‌ها	سرمایه‌گذاران	سرمایه‌گذاران	سوگیری	سوگیری	سوگیری اعتماد	سوگیری نژاد
۱	۰	۲	۱	۱	۴	۰	۰	۲	۲	۰	۱	۱	۰	۵۰,۰۰۰,۰۰۰	۱	۰	۱	۱	۰	۱	۱	۰	۱	۱	
۲	۱	۵	۰	۴	۴	۳	۲۰	۱۵	۷	۰	۰	۰	۱	۱۰۰,۰۰۰	۰	۱	۱	۰	۱	۱	۰	۱	۰	۱	۰
۳	۱	۲	۱	۳	۴	۰	۰	۰	۴	۰	۰	۰	۰	۲۰۰,۰۰۰	۰	۱	۰	۰	۰	۱	۰	۱	۱	۰	۰
۴	۱	۳	۰	۱	۳	۰	۰	۰	۲	۰	۰	۰	۱	۱,۰۰۰,۰۰۰	۰	۰	۱	۱	۱	۱	۱	۱	۱	۱	۰
۵	۱	۴	۰	۰	۴	۵	۷	۳	۱۰۰	۲	۰	۰	۱	۲۰۰,۰۰۰,۰۰۰	۱	۱	۱	۱	۱	۱	۰	۱	۰	۱	۱
۶	۱	۴	۰	۱	۴	۰	۵	۵	۲	۲	۰	۱	۱	۱۱,۰۰۰,۰۰۰	۱	۱	۰	۱	۱	۱	۱	۰	۱	۰	۰
۷	۱	۴	۰	۱	۴	۶	۵	۴	۰	۰	۰	۱	۱	۰	۱	۱	۰	۰	۱	۰	۱	۱	۱	۱	۰
۸	۱	۲	۱	۲	۳	۱	۴	۳	۵	۰	۰	۰	۱	۷,۰۰۰,۰۰۰	۰	۱	۱	۱	۱	۱	۱	۰	۱	۰	۱
۹	۱	۴	۰	۲	۴	۵	۰	۰	۲	۰	۰	۰	۰	۵,۰۰۰,۰۰۰	۰	۱	۱	۱	۱	۱	۰	۰	۱	۰	۰
۱۰	۱	۱	۱	۱	۴	۲	۰	۲	۵	۰	۰	۱	۱	۵,۰۰۰,۰۰۰	۰	۱	۱	۰	۱	۱	۱	۱	۱	۱	۰
۱۱	۰	۵	۱	۳	۴	۶	۱۲	۰	۲	۰	۰	۰	۱	۰	۱	۱	۰	۰	۱	۱	۱	۱	۱	۱	۱
۱۲	۰	۵	۰	۱	۳	۰	۷	۳	۲	۰	۰	۰	۱	۲۵,۰۰۰,۰۰۰	۰	۱	۱	۰	۱	۰	۱	۱	۱	۱	۰
۱۳	۰	۳	۰	۱	۴	۰	۵	۳	۱۵	۰	۰	۱	۱	۵۰,۰۰۰,۰۰۰	۰	۱	۰	۱	۱	۱	۱	۰	۰	۰	۰
۱۴	۱	۵	۰	۱	۴	۳	۰	۳	۴	۰	۰	۱	۱	۳,۰۰۰,۰۰۰	۱	۱	۰	۱	۱	۱	۱	۱	۰	۱	۰
۱۵	۱	۵	۱	۴	۵	۵	۰	۱۲	۱۲	۰	۱	۱	۱	۱۰,۰۰۰,۰۰۰	۰	۰	۱	۰	۰	۱	۱	۱	۱	۰	۱
۱۶	۱	۴	۱	۱	۴	۱	۰	۲	۱۲	۰	۱	۱	۱	۵۰,۰۰۰,۰۰۰	۱	۱	۰	۱	۰	۰	۱	۱	۰	۰	۰
۱۷	۰	۸	۱	۲	۴	۰	۰	۰	۵	۱	۰	۰	۰	۵۰۰,۰۰۰	۰	۱	۰	۱	۱	۱	۱	۱	۰	۰	۰
۱۸	۱	۶	۰	۱	۳	۴	۸	۰	۵	۲	۰	۱	۱	۱۰۰,۰۰۰,۰۰۰	۰	۰	۰	۱	۱	۰	۱	۱	۱	۱	۰
۱۹	۱	۴	۰	۴	۴	۵	۲۰	۱۲	۸	۰	۰	۱	۱	۱۵۰,۰۰۰,۰۰۰	۰	۱	۱	۱	۰	۱	۱	۰	۱	۰	۰
۲۰	۱	۴	۰	۰	۴	۳	۱	۲	۲۰	۰	۱	۱	۱	۱۵۰,۰۰۰,۰۰۰	۱	۱	۰	۱	۱	۱	۰	۰	۱	۰	۰
۲۱	۱	۳	۰	۲	۳	۵	۵	۲	۱۲	۰	۰	۰	۱	۱۰,۰۰۰,۰۰۰	۰	۱	۰	۰	۰	۰	۱	۰	۰	۱	۱

