

ارائه و پیاده‌سازی مدل پیشگویی مبتنی بر هوش کسب‌وکار جهت تصمیم‌گیری تشخیص‌گرا: مطالعه موردی بنگاه‌های سرمایه‌گذاری جمعی

محمد رضا شیخ عطار^{۱*}، حمزه خواستار^۲، رضا یوسفی زنونز^۳

- ۱- دانشجوی دکتری، دانشکده مدیریت و حسابداری، دانشگاه شهید بهشتی، تهران، ایران
- ۲- استادیار، گروه مدیریت کسب‌وکار، دانشکده مدیریت، دانشگاه خوارزمی، تهران، ایران
- ۳- استادیار، گروه مدیریت فناوری اطلاعات، دانشکده مدیریت، دانشگاه خوارزمی، تهران، ایران

پذیرش: ۱۳۹۶/۱۲/۸

دریافت: ۱۳۹۵/۱۱/۲۰

چکیده

بسترهای سرمایه‌گذاری جمعی به سایت‌هایی تبدیل شده‌اند که مردم می‌توانند هم از ایده‌های نو حمایت مالی انجام دهند و هم برای محصولات یا خدمات خود به دنبال جذب سرمایه باشند. در سال‌های اخیر، تجربیات موفق بسیاری در خصوص این نوع جذب سرمایه‌گذاری گزارش شده و افراد بسیاری جهت جلب سرمایه‌گذاری به این‌گونه سایت‌ها مراجعه کرده‌اند؛ اما علی‌رغم رشد چشم‌گیر استقبال افراد در نقش سرمایه‌گذار یا در نقش جذب‌کننده سرمایه، میزان درصد پروژه‌های موفق به جلب کامل سرمایه، رشد کاهنده داشته است. این چالش به دلیل عدم آگاهی و تجربه کافی از نحوه راه‌اندازی پروژه‌ها در سایت‌های سرمایه‌گذاری جمعی توسط افراد ایجاد شده است. به این منظور، در پژوهش کنونی با استفاده از ویژگی‌های هوش تجاری، ابتدا یک مدل جامع برای مسئله پیشگویی در کسب‌وکار هدف بر اساس شاخص‌های کلیدی عملکرد ارائه شده است. سپس مطابق نیاز استخراج‌شده از مدل، مجموعه بزرگی از داده‌ها از سایت کیک استارتر شامل سوابق پروژه، سوابق کاربر، داده زمانی و اطلاعات کاربران پروژه‌ها در رسانه‌های اجتماعی گردآوری شده و در مرحله بعد با توجه به مدل استخراج‌شده با ابزار داده‌کاوی، پیاده‌سازی مدل برای پیشگویی و ارزیابی تعهد مالی پروژه‌های راه‌اندازی شده انجام

گرفته است. نتایج پژوهش نشان می‌دهد که مدل‌های پیش‌بینی می‌توانند به‌طور مؤثر موفقیت پروژه‌ها را پیش‌بینی کرده و با استفاده از شناسایی عوامل موفقیت پروژه‌ها، پیشنهادهایی را در راستای بالا بردن احتمال موفقیت پروژه‌ها ارائه کنند.

واژگان کلیدی: هوش کسب‌وکار؛ شاخص‌های کلیدی عملکرد؛ سرمایه‌گذاری جمعی؛ داده‌کاوی.

۱- مقدمه

کسب‌وکارهای فعال در زمینه سرمایه‌گذاری جمعی^۱ یکی از گونه‌های جدید کسب‌وکار هستند که از طریق مفهوم وب ۲ بستری مناسب برای پیوند سرمایه‌گذاران و صاحبان ایده ایجاد می‌کنند. مدل عملیاتی این سایت‌های سرمایه‌گذاری جمعی می‌تواند مبتنی بر پاداش (مانند کیک استارتر^۲ و ایندیگوگو^۳)، مبتنی بر اهداء (مانند گیوفوروارد^۴) یا مبتنی بر ارائه سهام (مانند ارلی شیرز^۵) باشد. تمامی این بسترها با استفاده از مدل‌های ارائه خدمات سرمایه‌گذاری جمعی موفق شده‌اند از طریق میلیون‌ها حامی، میلیاردها دلار را برای راه‌اندازی پروژه‌ها جذب کنند.

هرچند نمونه‌های زیادی از این پروژه‌ها وجود دارند که توانسته‌اند با اقبال زیادی از سوی سرمایه‌گذاران روبرو شوند، اما میزان موفقیت پروژه‌ها در جذب سرمایه در سایت‌های سرمایه‌گذاری جمعی، روند نزولی داشته است. علاوه بر این مشکل، اطلاعات اندکی نیز در مورد خصوصیات پروژه‌های موفق در بستر سرمایه‌گذاری جمعی وجود داد. از این‌رو، در راستای پاسخ به این خلأ اجرایی و پژوهشی، در مقاله حاضر با استفاده از روش‌های گام‌به‌گام مدلی جهت پیاده‌سازی الگوی پیش‌بینی کسب‌وکار هدف ارائه شده است. همچنین متغیرهای لازم جهت پیاده‌سازی الگو شناسایی شده و پس از جمع‌آوری داده‌ها، با استفاده از مدل‌سازی سناریوهای پیش‌بینی پیاده‌سازی شده‌اند.

1. Crowd Funding
2. Kickstarter
3. Indiegogo
4. GiveForward
5. EarlyShares

در ادامه مقاله، ابتدا در بخش پیشینه نظری، اصطلاحات علمی به‌کاررفته در مقاله مختصراً بیان شده، سپس در پیشینه عملی کارهای انجام‌شده در این زمینه ارائه شده است. در بخش روش تحقیق چارچوب پیشنهادی پیاده‌سازی هوش کسب‌وکار هدف ارائه شده و به پیاده‌سازی چارچوب پیشنهادی هوش کسب‌وکار مطابق الگو پرداخته شده است. این چارچوب شامل شناخت کسب‌وکار، فرمول‌بندی شاخص کلیدی عملکرد، شناسایی و معرفی مشخصه‌های لازم، تعیین نوع مشخصه‌ها، ساخت مدل و تجزیه و تحلیل بهترین تابع پیشگویی‌کننده است.

۲- پیشینه پژوهش

۲-۱- هوش کسب‌وکار

هوش کسب‌وکار فرآیندی است که در آن داده‌ها به اطلاعاتی تبدیل می‌شود که برای تصمیم‌سازی موردنیاز است. تعاریف متعددی در مبنای نظری برای هوش تجاری ارائه شده است. سیستم‌های هوش تجاری قابلیت را ایجاد می‌کنند که به‌واسطه آن می‌توان اطلاعات تجاری را تجزیه و تحلیل کرد. این تجزیه و تحلیل اطلاعات موجب تقویت و پشتیبانی از فرآیند تصمیم‌سازی در محدوده وسیعی از فعالیت‌های کسب‌وکار می‌شود [۱]. این سیستم‌ها موجب ایجاد سیستم‌های توصیه‌گر به‌منظور ارائه پیشنهاد به مشتریان و ارائه خدمات می‌شوند [۲].

هوش تجاری به‌مثابه فرآیندی مداوم است که به سازمان‌ها کمک می‌کند به جمع‌آوری و ثبت داده پرداخته، داده‌ها را تجزیه و تحلیل و اطلاعات و دانش به‌دست‌آمده را جهت ارتقا کارکرد سازمانی در فرآیند تصمیم‌سازی به کار گیرند [۳]. از این دیدگاه، هوش تجاری به‌صورت یک چرخه شامل ثبت، پردازش و واکنش بیان می‌شود. در مرحله ثبت، حلقه هوش کسب‌وکار از طریق گوش دادن دقیق به محیط (ثبت اطلاعات) آغاز می‌شود. در مرحله پردازش، بعد از آنکه داده‌ها ثبت شدند، لازم است پردازش انجام شود. پردازش داده‌های جمع‌آوری‌شده روند را آشکار و اطلاعات باارزشی را ارائه می‌کند. نهایتاً در مرحله واکنش، در ادامه نتایجی که به‌وسیله تحلیل داده‌ها حاصل می‌شود، شرکت می‌تواند واکنش لازم را در سه سطح عملیاتی، تاکتیکی و استراتژیک طراحی و اجرا کند. در نتیجه، محیط تغییرات شرکت را ارزیابی کرده، موجب ایجاد سیگنال‌های جدید برای چرخه هوش کسب‌وکار می‌شود [۴].

۲-۲- تأمین مالی جمعی

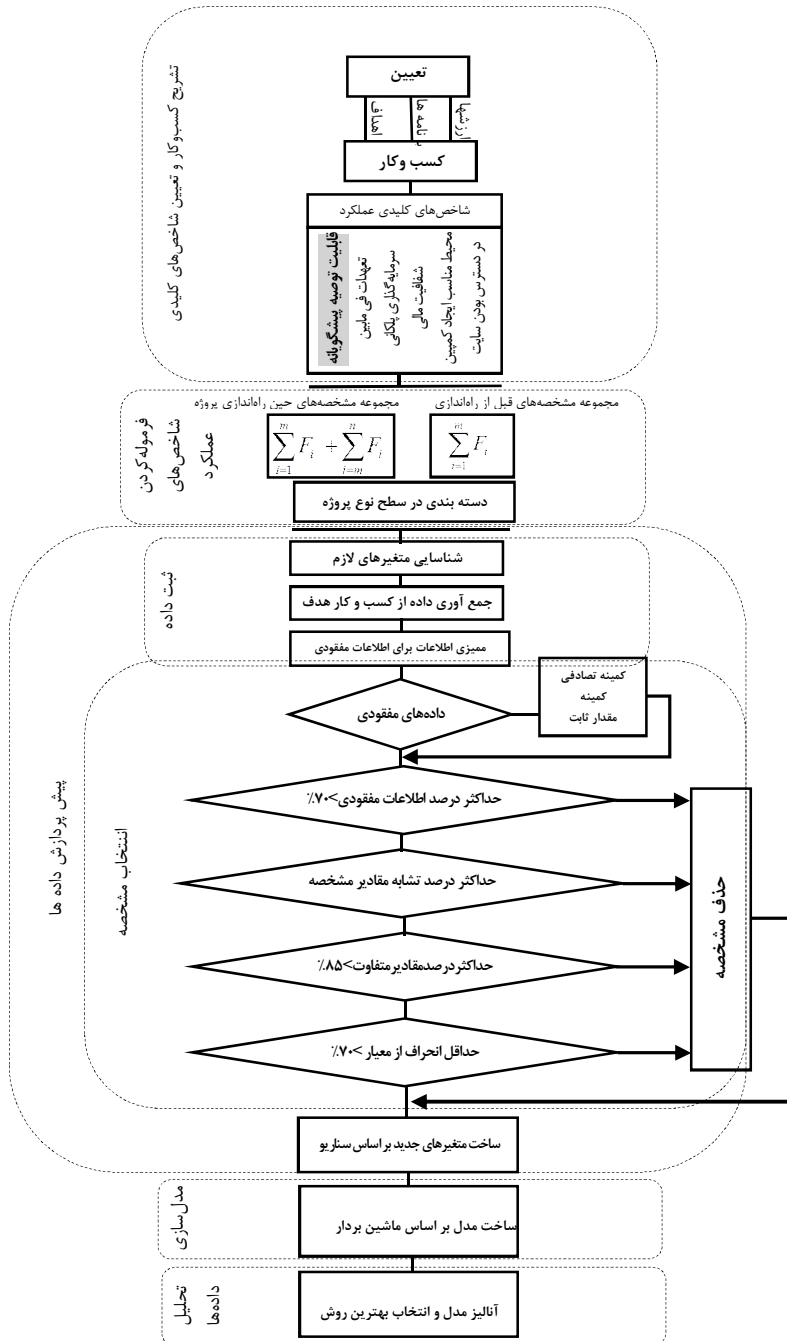
تأمین مالی جمعی، گونه‌ای از جمع‌سپاری در حوزه تأمین مالی کسب‌وکارهای نوپا و ایده‌های نوآورانه است. این ایده نیز مبتنی بر فراخوان باز و جمعیت بوده^۱ و در آن «خرد جمعی» است که در تصمیم‌گیری‌ها عمل می‌کند [۵]. تأمین و توزیع بودجه در تأمین مالی جمعی مبتنی بر تمرکززدایی است. در نظام مالی جمعی، تأمین و مسیریابی منابع مالی به‌سوی فرصت‌های سرمایه‌گذاری و کارآفرینی، از «مدل شبکه‌ای و توزیع‌شده» تبعیت می‌کند. با این توصیف که کارآفرینان و سرمایه‌گذاران (مشارکت‌کنندگان) در یک بستر مشارکتی که نوع تکامل‌یافته‌ای از شبکه‌های اجتماعی است، به تعامل با یکدیگر می‌پردازند. در ابتدا، آمارها از موفقیت چشم‌گیر پروژه‌های راه‌اندازی شده در جلب سرمایه‌گذاری خبر می‌دادند، اما با رشد پروژه‌های راه‌اندازی شده و سرمایه‌گذاران، روند موفقیت پروژه‌ها رو به کاهش گذاشته است و این به دلیل نبود تجربه در نوع ارائه ایده‌ها و آمادگی ناکافی ارائه‌کننده‌های پروژه بوده است. زو، یانگ و رائو برای حل این مشکل، ابتدا یک مجموعه داده بزرگ از یک‌استارتر را جمع‌آوری کردند که متشکل از پروفایل‌های پروژه (مربوط به پروفایل کاربر)، داده زمانی پروژه و اطلاعات رسانه‌های اجتماعی کاربران است. در مرحله بعد ویژگی‌های پروژه‌های موفق و رفتار کاربران را مورد تجزیه و تحلیل قرار دادند. در مرحله بعد، مدلی آماری را برای پیش‌بینی موفقیت پروژه‌ها برای جذب سرمایه و تعهد مالی کاربران ارائه کردند. در این راستا، ارتباط کلیه متغیرها را با موفقیت یا عدم موفقیت تأمین مالی پروژه‌ها استخراج نموده و در آخر نیز به توسعه و ارزیابی مدل پرداختند [۶].

۳- چارچوب پیشنهادی پیاده‌سازی هوش کسب‌وکار در کسب‌وکار

هدف

چارچوب پیشنهادی پیاده‌سازی هوش کسب‌وکار در کسب‌وکار هدف شامل ۵ قسمت کلی شناخت کسب‌کار، فرمول‌بندی شاخص کلیدی عملکرد، جمع‌آوری داده، مدل‌سازی و ارزیابی است که نحوه پیاده‌سازی و جزئیات آن در شکل زیر تشریح شده است:

1. Open call and crowd



شکل ۱ چارچوب پیشنهادی پیاده‌سازی هوش کسب و کار

۳-۱- شناخت کسب‌وکار

در این مرحله از فرآیند هوش کسب‌وکار است که مدیران می‌توانند از طریق آن مشخص‌کننده با چه فعالیت‌هایی ارزش‌ها به اهداف اصلی اضافه شده است و چرا و چگونه کارایی این فعالیت‌ها اندازه‌گیری می‌شود. در این راستا، چارچوب‌های مختلفی به جهت ساختن یک مدل کسب‌وکار بر اساس استراتژی شرکت وجود دارد که در آن -در انتها- معیارهای کارایی می‌تواند فرموله شود [۷]. در این تحقیق، مطابق اهداف کسب‌وکاری که در سایت کیک‌استارتر قرار گرفته است و مدل‌های کسب‌وکار ارائه‌شده توسط ایجادکننده‌های سایت، چند شاخص کلیدی عملکرد به‌عنوان نمونه برای ورود به مرحله بعد تدوین شد که شامل موارد زیر است: قابلیت توصیه توسط سایت جهت راه‌اندازی پروژه‌ها، تعهدات فی‌مابین، سرمایه‌گذاری پلکانی، شفافیت مالی، محیط مناسب ایجاد پویش، در دسترس بودن سایت و امنیت سرمایه‌گذاری [۸].

۳-۲- فرمول‌بندی شاخص کلیدی عملکرد انتخابی

در این قسمت، دو سناریو برای فرمول‌بندی شاخص‌های کلیدی عملکرد در نظر گرفته شده است. اول، پیشگویی قبل از راه‌اندازی در سایت که از مجموع مشخصه‌هایی که برای پیشگویی لازم است، تعداد m مشخصه از کل n مشخصه انتخاب شده است. این مشخصه‌ها شامل وضعیت، واحد پول، پیشنهاد کیک‌استارتر، تعداد ویدئو، تصویر، عکس و لینک فیس‌بوک است.

سناریو دوم، پیشگویی در حال اجرای پروژه است که شامل $m-n$ متغیر باقیمانده شامل هدف، جذب سرمایه، مدت پویش^۱، تعداد به‌روز شدن‌ها در خلال پروژه و غیره است. البته همان‌طور که در شکل مدل هم قابل مشاهده است، مشخصه‌های پیش از راه‌اندازی پروژه در سناریوی دوم شرکت داده شده‌اند.

۳-۳- شناسایی و معرفی مشخصه‌های لازم

با استفاده از فرمول‌بندی شاخص‌های کلیدی عملکرد و سناریو در نظر گرفته‌شده در این مرحله، کلیه مشخصه‌های لازم از سایت سرمایه‌گذاری جمعی هدف شناسایی شد که در جدول ۱ ارائه شده است. گفتنی است مشخصه‌هایی که بعد از

اتمام مرحله جذب سرمایه کامل می‌گردند، مطابق سناریو فاقد ارزش بوده و کنار گذاشته شده است.

جدول ۱ مشخصه‌ها^۱ و مقادیر نمونه

نام مشخصه	مقادیر نمونه
متوسط حاصل	۵۷
ایجاد شده به وسیله خلق	۷
پشتیبانی از خلافت	۸
مکان	اورلاندو
نام خلق	ردیفک
تعداد کلمات پوشش	۸۳۳
تعداد کلمات مخفی‌آمیز	۲۲
اینکه قس بوک	خیر
دیدگاهها ^۲ بر خلق پوشش	۰
بروز شدن‌ها در خلال پوشش	۱
حاصل	۴
عکس	۰
تصویر	۰
تعداد ویدیو	۰
ویدیو تأمین مالی	خیر
روز پایان	سپتامبر
تاریخ پایان	۲۰۰۹-۱۰-۲۳
روز شروع	چهارشنبه
تاریخ شروع	۲۰۰۹-۰۶-۰۳
مدت پوشش	۱ روز
پیشینه کیک‌استارتز	خیر
واحد پول	دلار
درصد جذب سرمایه	۹ درصد
جذب سرمایه	۲۳۰
هاف	۲۴۹
وضعیت	شکست
نام پوشش	بازی و پیشگویی از راه
شماره پروژه	۶۱۹۸

۳-۴- جمع‌آوری داده

یکی از معتبرترین سایت‌های سرمایه‌گذاری جمعی سایت کیک‌استارتز است. در این بستر کاربران اقدام به ایجاد و جلب سرمایه‌گذاری و حمایت مالی از پروژه‌ها می‌نمایند. از ماه می ۲۰۱۷ این سایت دارای رتبه ۵۰۶ بین سایت‌های پربازدید دنیا بوده است [۹]. هدف از جمع‌آوری داده‌های کیک‌استارتز، جمع‌آوری سوابق پروژه‌ها مطابق مشخصه‌های استخراج‌شده از سناریو در مرحله قبل است. خوشبختانه کیک‌استارتز تمامی سوابق پروژه‌ها را در یک آدرس جداگانه^۳ قرار داده است. روش جمع‌آوری داده به این صورت بود که هر هر پروژه توسط نرم‌افزار خاص باز شده و اطلاعات مربوط به مشخصه‌ها از آن خوانده شده و به صورت فرمت CSV در یک فایل ذخیره می‌شد. در نهایت، اطلاعات مربوط به ۸۰۰۱ پروژه بین سال‌های ۲۰۱۵ تا ۲۰۱۶ که شامل نام

1. Fitures
2. Comment
3. <http://www.kicktraq.com/archive/>

پویش، وضعیت، هدف، جذب سرمایه، واحد پول، پیشنهاد کیک استارتر و غیره بود استخراج شد.

۳-۵- تعیین نقش و نوع مشخصه‌ها

با توجه به سناریو پیشگویی که شامل پیشگویی موفقیت و عدم موفقیت پروژه‌ها است، مشخصه وضعیت پروژه‌ها (موفق، ناموفق، متوقف‌شده) به‌عنوان متغیر وابسته یا خروجی و بقیه متغیرها به‌عنوان متغیرهای مستقل یا ورودی در نظر گرفته شد (جدول ۲).

جدول ۲ نمایش متغیرهای مستقل و وابسته

<p>متغیر خروجی وضعیت حین راه‌اندازی</p>	$Y_1 = a_1X_1 + a_2X_2 + \dots + a_{m-1}X_{m-1} + a_mX_m + a_{m+1}X_{m+1} + a_{m+2}X_{m+2} + \dots + a_{n-1}X_{n-1} + a_nX_n$
	<p>مشخصه‌های ورودی پیش از راه‌اندازی پروژه</p>
	<p>مشخصه‌های ورودی حین راه‌اندازی پروژه</p>
<p>متغیر خروجی وضعیت قبل از راه‌اندازی</p>	$Y_2 = a_1X_1 + a_2X_2 + \dots + a_{m-1}X_{m-1} + a_mX_m$
	<p>مشخصه‌های ورودی پیش از راه‌اندازی پروژه</p>

۳-۶- داده‌های مفقودی

به دلیل نبود اطلاعات در بعضی مشخصه‌های پروژه، با اطلاعات مفقودی مواجه بودیم. با این اطلاعات مفقودی به دو صورت برخورد شد: آن دسته از مشخصه‌هایی که دارای اطلاعات مفقودی کم بودند، متوسط تصادفی برای اطلاعات مفقودی آن‌ها در نظر گرفته شد و برای مشخصه‌هایی که دارای داده‌های مفقودی زیادی بودند، مقادیر ثابت در نظر گرفته شد.

۳-۷- فرآیند انتخاب مشخصه (غربالگری مشخصه‌ها)

برای مقادیر مشخصه‌ها آستانه‌هایی تعریف شد که شامل ۱- حداکثر درصد اطلاعات مفقودی: ردیف‌هایی که مقادیر مفقودی زیاد دارند، حذف شوند، ۲- حداکثر درصد تشابه مقادیر مشخصه: ردیف‌هایی که رکوردهای زیادی از آن‌ها در دسته‌های مشابه قرار گرفته‌اند و ۳- حداکثر درصد مقادیر متفاوت مشخصه: ردیف‌هایی را غربال می‌کند که مقدار زیادی دسته‌های جزئی و متفاوت با سایر دسته‌ها در این ردیف‌ها وجود دارند. مقادیر حدی در جدول ۳ نشان داده شده است.

جدول ۳ حذف مشخصه‌های غیرعادی

مقادیر آستانه	ویژگی حذف مشخصه‌ها
۷۰ درصد	حداکثر درصد اطلاعات مفقودی
۸۸ درصد	حداکثر درصد تشابه مقادیر مشخصه
۸۵ درصد	حداکثر درصد مقادیر متفاوت مشخصه
۰/۱	حداقل ضریب تنوع
۰/۱	حداقل انحراف از معیار

۳-۸- ساخت مدل بر اساس ماشین بردار پشتیبان

در این مقاله برای مدل‌سازی از یک ماشین بردار پشتیبانی که بر اساس شبکه عصبی چندگانه عمل می‌کند، استفاده شده است. برای استفاده از ماشین بردار پشتیبانی لازم است که تابع کرنل مشخص شود [۱۰]. در اینجا، چهار تابع کرنل به نام‌های RBF، Polynomial، Linear و Sigmoid داریم. نحوه اندازه‌گیری خطای تابع ماشین بردار پشتیبان با کرنل‌های مختلف به این صورت است که در اینجا ما تعدادی داده آموزشی داریم، یک مجموعه از نقاط به شکل:

$$D = \{(x_i, y_i) | x_i \in R^p, y_i \in \{1, -1\}\} \quad i = 1, \dots, n$$

که y_i می‌تواند ۱ یا -۱ باشد و مشخص می‌کند نقطه x_i متعلق به چه کلاسی است. هر x_i یک بردار p بعدی است. می‌خواهیم جداکننده بیشینه حاشیه‌ای پیدا کنیم که نقاطی

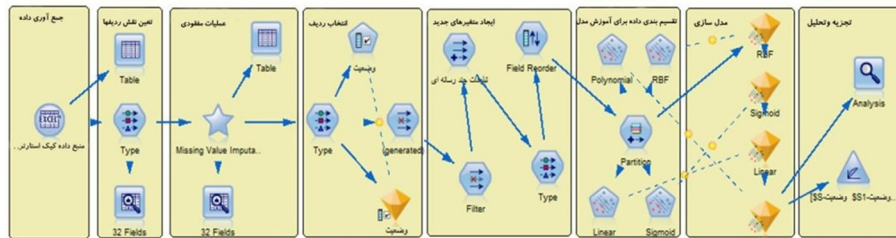
که y_i آن‌ها برابر ۱ است از نقاطی که y_i آن‌ها -۱ است، جدا کند. هر ابر صفحه می‌تواند به صورت یک مجموعه از نقاط x که در رابطه زیر صدق می‌کنند، نوشته شود.

$$w_1x_1 + w_2x_2 + \dots + w_nx_n + b = 0 \quad \text{و} \quad \langle w, x \rangle + b = 0$$

می‌خواهیم مقادیر W و b را به گونه‌ای پیدا کنیم که نمونه‌های آموزشی را به دقت دسته‌بندی کند؛ با این فرض که داده‌ها به صورت خطی جدا پذیر باشند، فاصله بین ابر صفحه‌های موازی یعنی پهناي حاشیه را حداکثر نماید. این ابر صفحه‌های موازی می‌توانند با معادله‌های ذیل توصیف شوند:

$$W.X + b \leq 1 \quad \forall x_i \quad \text{و} \quad W.X + b \geq 1 \quad \forall x_i$$

با استفاده از روابط هندسی پهناي حاشیه برابر است با $\frac{2}{\|w\|}$. بنابراین می‌خواهیم w را کمینه کنیم. همچنین می‌خواهیم از قرارگرفتن نقاط در حاشیه جلوگیری کنیم. محدودیت زیر را با ترکیب دو رابطه بالا اضافه می‌کنیم: $y_i (w \cdot x_i + b) \geq 1$. حل مسئله بهینه‌سازی $\frac{2}{\|w\|}$ مشکل است، زیرا به w بستگی دارد. خوشبختانه می‌توان معادله را با توجه به رابطه $\|w\| = 1/2\|w\|^2$ بدون تغییر راه‌حل، جایگزین کرد. حل این مسئله از طریق ضرایب لاگرانژ یا مسئله دوگان قابل حل است [۱۱]. در واقع، رویکرد ماشین بردار پشتیبانی به این صورت است که در مرحله آموزش، سعی دارد که مرز تصمیم‌گیری را به گونه‌ای انتخاب نماید که حداقل فاصله آن با هر یک از دسته‌های موردنظر را بیشینه کند تا خطا به حداقل برسد. این نوع انتخاب باعث می‌شود که تصمیم‌گیری ما در عمل، شرایط نوپزی را به خوبی تحمل نموده و همچنین پاسخ‌دهی مناسبی داشته باشد [۱۱]. میزان خطای ماشین بردار پشتیبان با توابع کرنل مختلف در جداول ۴ تا ۷ از سناریو اول و جداول ۸ تا ۱۱ از سناریو دوم دیده می‌شود. این میزان خطا مبنای انتخاب بهترین تابع کرنل برای ماشین بردار تصمیم بوده است. برای مقایسه عملکرد پیشگویی این توابع هرکدام از این توابع را اجرا می‌کنیم و هرکدام را که عملکرد بهتری داشت و دارای خطای کمتری بود، برای مدل ارائه شده در شکل ۲ انتخاب می‌کنیم.



شکل ۲ نمای طرح مدل در نرم‌افزار

۹-۳- تجزیه و تحلیل مدل و انتخاب بهترین تابع پیشگویی‌کننده

در این بخش، خروجی‌های سه سناریوی مدل‌سازی شده مطابق چارچوب پیشنهادی هوش کسب‌وکار ارائه شده و خروجی‌ها مورد بررسی و تحلیل قرار می‌گیرند:

سناریوی اول: مدل پیشگویی قبل از راه‌اندازی پروژه

نتایج ماتریس درهم‌ریختگی^۱ در جداول ۴ الی ۷ ملاحظه می‌شوند. در اینجا، داده‌ها به ۷۰ درصد داده‌های آموزشی و ۳۰ درصد داده‌های آزمون تقسیم شدند. همان‌طور که از جدول ۶ درهم‌ریختگی دیده می‌شود، تابع خطی با ۶۲/۴۵ درصد موفقیت در پیشگویی، عملکرد بهتری نسبت به بقیه توابع داشته است؛ البته بقیه توابع نیز عملکرد نزدیکی با تابع خطی داشتند. در شکل ۴ می‌توان عملکرد توابع را مشاهده کرد. همان‌طور که در نمودار ملاحظه می‌شود توابع RBF، زیگموئید و چندجمله‌ای عملکرد نزدیکی در پیشگویی موفقیت پروژه‌ها قبل از راه‌اندازی نسبت به یکدیگر دارند. نکته مهمی که در تحلیل نتایج می‌توان مشاهده کرد این است که با آنکه تابع چندجمله‌ای در قسمت آموزش بهترین عملکرد را داشته (۶۴/۸۵ درصد) اما در آزمون و پیشگویی نتوانسته عملکرد خوبی داشته باشد (۶۲/۰۱ درصد) و این به دلیل آن است که درجه آزادی مدل بسیار بیشتر از درجه آزادی واقعی انتخاب شده و در نتیجه، اگرچه مدل روی داده استفاده‌شده برای یادگیری بسیار خوب نتیجه می‌دهد، اما بر روی داده جدید دارای خطا است. با توجه به اینکه تابع خطی، هرچند ناچیز، عملکرد بهتری نسبت به بقیه داشت، به‌عنوان بهترین تابع انتخاب شد. به این منظور، از نقطه نظر تابع خطی مهم‌ترین مشخصه‌ها

1. Confusion matrix

مطابق شکل ۳ استخراج شد. همان‌طور که در شکل دیده می‌شود، مشخصه پیشنهاد کیک‌استارتر دارای بالاترین رتبه در پیشگویی و بعد از آن به ترتیب، واحد پول و حمایت از خلاقان در رتبه‌های بعدی قرار دارند.

جدول ۴ عملکرد تابع RBF در پیشگویی قبل از راه‌اندازی

درصد	آزمون	درصد	آموزش	قسمت‌بندی
۶۲/۳۸	۹۹۶	۶۳/۵	۴۰۶۴	صحیح
۳۷/۶۲	۵۹۹	۳۶/۵۵	۲۳۴۱	خطا
	۱۵۹۵		۶۴۰۵	مجموع

جدول ۵ عملکرد تابع زیگموئید در پیشگویی قبل از راه‌اندازی

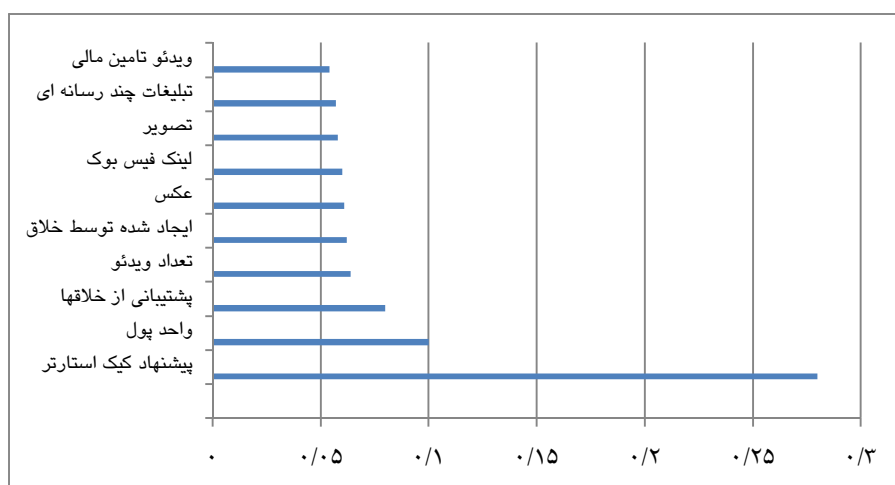
درصد	آزمون	درصد	آموزش	قسمت‌بندی
۵۹/۳۷	۹۴۳	۶۱/۳۷	۲۹۳۵	صحیح
۴۰/۶۳	۶۵۲	۳۸/۶۳	۲۴۷۰	خطا
	۱۵۹۵		۶۴۰۵	مجموع

جدول ۶ عملکرد تابع خطی در پیشگویی قبل از راه‌اندازی

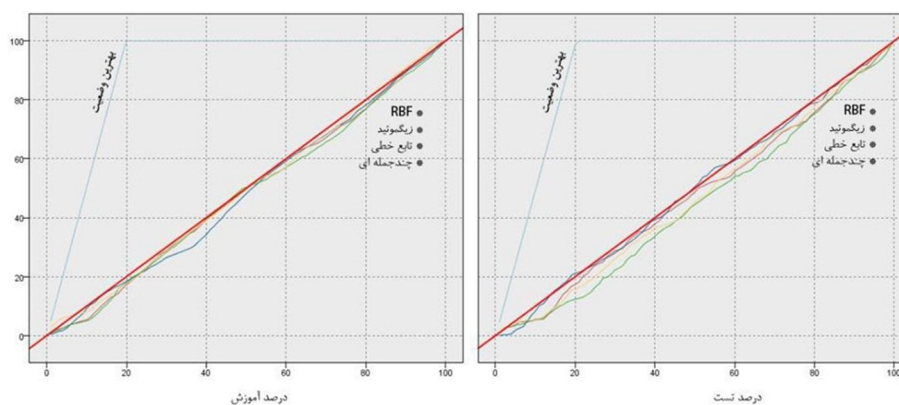
درصد	آزمون	درصد	آموزش	قسمت‌بندی
۶۲/۴۵	۹۹۶	۶۳/۲۶	۴۰۵۲	صحیح
۳۷/۵۵	۵۹۹	۳۶/۷۴	۲۳۵۳	خطا
	۱۵۹۵		۶۴۰۵	مجموع

جدول ۷ عملکرد تابع چندجمله‌ای در پیشگویی قبل از راه‌اندازی

درصد	آزمون	درصد	آموزش	قسمت بندی
۶۲/۰۱	۹۹۶	۶۴/۸۵	۴۱۱۳	صحیح
۳۷/۹۹	۵۹۹	۳۴/۱۵	۲۲۹۲	خطا
	۱۵۹۵		۶۴۰۵	مجموع



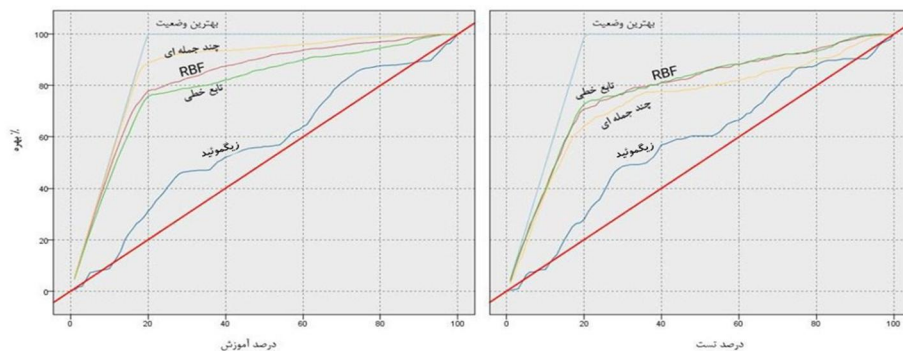
شکل ۳ ترتیب اهمیت مشخصه‌های پیشگویی کننده قبل از راه‌اندازی پروژه



شکل ۴ نمودار مقایسه عملکرد توابع ماشین بردار تصمیم در پیشگویی قبل از راه‌اندازی

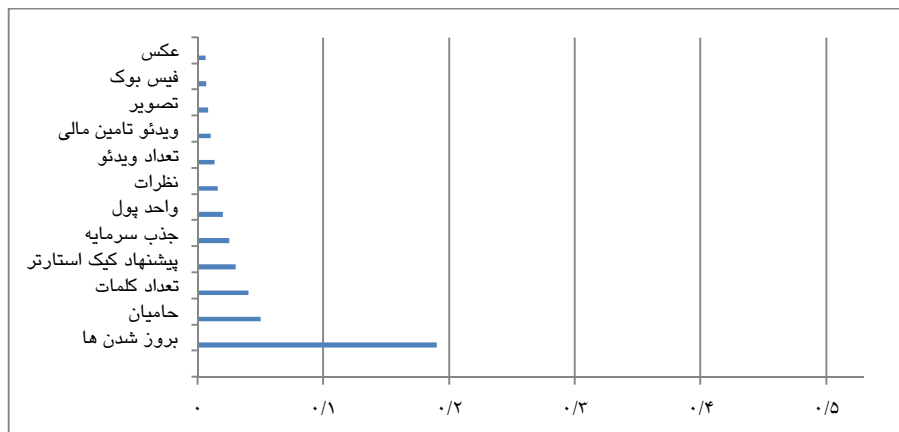
سناریوی دوم: مدل پیشگویی حین راه‌اندازی پروژه

در اینجا، نتایج ماتریس درهم‌ریختگی حاصل از عملکرد ماشین بردار تصمیم در جداول ۸ تا ۱۱ مشاهده می‌شود. در اینجا نیز ما داده‌ها را به ۷۰ درصد داده‌های آموزشی و ۳۰ درصد داده‌های آزمون تقسیم کردیم. همان‌طور که از جدول درهم‌ریختگی قابل مشاهده است، تابع خطی با $77/37$ درصد موفقیت در پیشگویی، عملکرد بهتری نسبت به بقیه توابع داشته است. در شکل ۵ می‌توان عملکرد توابع را مشاهده کرد. همان‌طور که در نمودار دید می‌شود، ترتیب عملکرد توابع در پیشگویی موفقیت یا عدم موفقیت پروژه‌ها در جذب سرمایه به ترتیب خطی (با $77/37$ درصد)، RBF ($75/74$ درصد)، چندجمله‌ای ($73/92$ درصد)، و زیگمویید ($59/56$ درصد) است. در اینجا نیز نکته مهمی که در تحلیل نتایج می‌توان مشاهده کرد، این است که با آنکه تابع چندجمله‌ای در قسمت آموزش بهترین عملکرد را داشته ($88/59$ درصد) اما در آزمون و پیشگویی نتوانسته عملکرد خوبی داشته باشد ($73/92$ درصد) و این به دلیل بیش‌برازش^۱ در یادگیری آماری است. همچنین مطابق عملکرد تابع خطی مهم‌ترین شاخصه‌های پیشگویی‌کننده مطابق شکل ۶ ارائه شده است. همان‌طور که در شکل دیده می‌شود، مشخصه مدت پویش دارای بالاترین رتبه در پیشگویی و بعد از آن به ترتیب، به روز شدن در خلال پویش و حامیان در رتبه‌های بعدی قرار دارند.



شکل ۵ نمودار مقایسه عملکرد توابع ماشین بردار تصمیم در پیشگویی حین راه‌اندازی

1. Overfitting



شکل ۶ ترتیب اهمیت مشخصه‌های پیشگویی‌کننده حین راه‌اندازی پروژه

جدول ۸- عملکرد تابع RBF در پیشگویی حین راه‌اندازی

قسمت‌بندی	آموزش	درصد	آزمون	درصد
صحیح	۱۴۳۵	۳۵	۲۰۸۱	۷۴/۷۵
خطا	۲۶۲۱	۶۵	۳۸۷	۲۶/۲۴
مجموع	۴۰۵۶		۵۹۵۱	

جدول ۹ عملکرد تابع زیگموئید در پیشگویی حین راه‌اندازی

قسمت‌بندی	آموزش	درصد	آزمون	درصد
صحیح	۳۱۸۸	۷۸/۶۱	۹۵۰	۵۶/۵۹
خطا	۸۶۸	۲۲/۳۸	۶۴۵	۴۴/۴۰
مجموع	۴۰۵۶		۵۹۵۱	

جدول ۱۰ عملکرد تابع خطی در پیشگویی حین راه‌اندازی

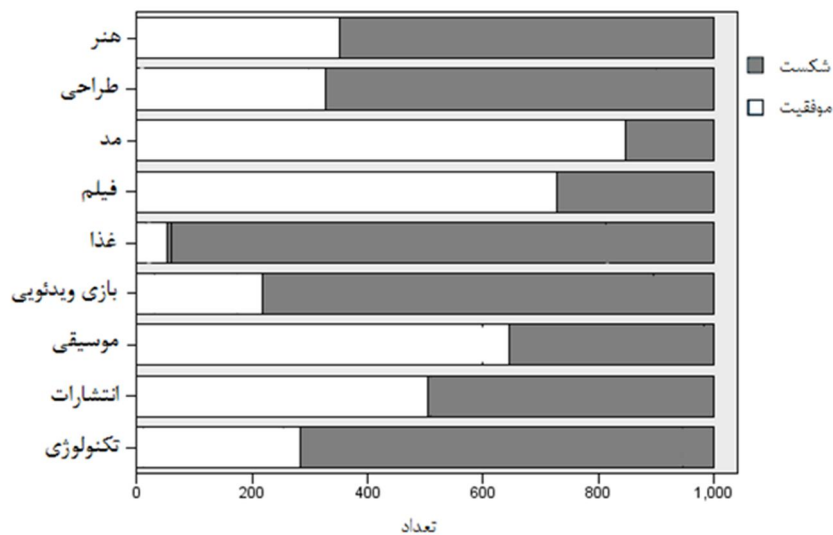
قسمت‌بندی	آموزش	درصد	آزمون	درصد
صحیح	۸۰۵	۳۱/۷۹	۲۳۴۱	۷۷/۳۷
خطا	۳۲۵۱	۶۹/۲۰	۳۶۱	۶۳/۲۲
مجموع	۴۰۵۶		۵۹۵۱	

جدول ۱۱ عملکرد تابع چندجمله‌ای در پیشگویی حین راه‌اندازی

قسمت بندی	آموزش	درصد	آزمون	درصد
صحیح	۵۶۷۴	۸۸/۵۹	۱۱۷۹	۷۳/۹۲
خطا	۷۳۱	۱۱/۴۱	۴۱۶	۲۶/۰۸
مجموع	۶۴۰۵		۱۵۹۵	

سناریوی سوم: پیشگویی بر مبنای نوع پروژه

در سناریوی سوم، اطلاعات مربوط به نوع پروژه‌ها استخراج شده است. از هر نوع پروژه راه‌اندازی شده حدود هزار رکورد، مجموعاً ۹۰۰۰ رکورد، استخراج شد تا تحلیل طبقه‌بندی نوع پروژه‌ها با در نظر گرفتن نوع پروژه به‌عنوان متغیر ورودی و وضعیت پروژه به‌عنوان متغیر خروجی، صورت پذیرد که نتیجه در شکل ۷ ارائه شده است. از طریق جمع‌آوری داده‌های طبقه‌بندی پروژه‌ها می‌توان به میزان موفقیت یا شکست، بر اساس نوع پروژه‌ها، دست یافت. همان‌طور که در شکل ۷ دیده می‌شود پروژه‌های مد ۸۰ درصد، فیلم ۷۸ درصد، موسیقی ۶۵ درصد، انتشارات ۵۰ درصد احتمال موفقیت در جذب سرمایه را از طریق تحلیل نوع پروژه دارند.



شکل ۷ نمودار پیشگویی بر مبنای نوع پروژه

۴- بحث و نتیجه‌گیری

کارهایی که در زمینه پیشگویی پروژه‌های موفق انجام شده، صرفاً متمرکز بر بهینه‌سازی مدل‌های پیشگویی جهت ارتقاء عملیات پیشگویی بوده است و به سناریوی پیشگویی پرداخته نشده است. هدف مقاله حاضر ارائه پیشنهادهایی مبتنی بر اطلاعات موجود از وضعیت پروژه‌ها از طریق پیشگویی موفقیت پروژه‌های راه‌اندازی شده با استفاده از سناریوسازی در هوش کسب‌وکار است. در این مقاله بحث شد که شناسایی اهداف و مشخص کردن سناریوهای مختلفی که بتواند این اهداف کسب‌وکار را پوشش دهد، تأثیر عمیقی بر روی مدل‌سازی و متعاقب آن بر روی تحلیل داده‌ها خواهد گذاشت. در واقع اینکه چه مشخصه‌هایی و چه داده‌هایی باید در هوشمند سازی کسب‌وکار رصد شوند، به این مسئله وابسته است که تا چه حد توانسته باشیم اهداف کسب‌وکار را شناسایی و سناریوهای مناسب آن را طراحی کرده باشیم. در ادامه بعد از شناسایی شاخص کلیدی عملکرد نمونه، سه سناریو برای فرموله کردن ارائه شد که سناریو اول، پیشگویی قبل از راه‌اندازی، سناریوی دوم پیشگویی حین راه‌اندازی و سناریوی سوم، دسته‌بندی در سطح تنوع پروژه بود. مطابق این سناریو مشخصه‌های موردنظر برای پوشش سناریو از سایت کیک‌استارتر استخراج و پیش‌پردازش و مدل‌سازی برای انتخاب بهترین مدل از لحاظ عملکرد پیشگویی، مطابق سناریو صورت گرفت. در این راستا، سناریو اول که پیشگویی قبل از راه‌اندازی پروژه بود، مشخصه پیشنهاد کیک‌استارتر (پیشنهادی است که سایت کیک استارتر به مشتریان خود قبل از راه‌اندازی پروژه در خصوص سرمایه‌گذاری روی پروژه‌ها می‌دهد) از اهمیت بالایی برخوردار است (شکل ۳). این جواب سؤال پژوهشی در خصوص وجود مشخصه‌های محیطی و تأثیر آن‌ها در پیشگویی بود. این مشخصه و انتخاب آن به‌عنوان یک مشخصه کلیدی توسط پیشگویی‌کننده نشان می‌دهد که قبل از سرمایه‌گذاری، کاربران ابتدا روی پروژه‌هایی سرمایه‌گذاری می‌کنند که سایت کیک‌استارتر آن‌ها را پیشنهاد داده باشد و این نشان از اعتماد مردم به این سایت دارد. عامل مهم پشتیبانی از خلاقان (سومین مشخصه مهم پیشگویی که نشان از حمایت خود خلاق از پروژه‌های دیگر دارد) و تعداد پروژه‌های راه‌اندازی شده (ششمین مشخصه مهم پیشگویی) نشان می‌دهد اگر کاربران در راه‌اندازی

پروژه‌ها دارای تجربه باشند، احتمال موفقیت تأمین سرمایه موردنظر توسط سرمایه‌گذاران بیشتر خواهد بود؛ بنابراین، سؤال پژوهش در خصوص تجربه و تأثیر آن در موفقیت پروژه‌ها پاسخ داده شد. همچنین از عوامل مهم در پیشگویی، تعداد ویدئو از پروژه (چهارمین مشخصه مهم پیشگویی)، وجود لینک فیس‌بوک برای تبلیغ (هشتمین مشخصه مهم پیشگویی) و تصویر از پروژه (نهمین مشخصه مهم پیشگویی) نشان می‌دهد که پروژه‌هایی که قبل از راه‌اندازی اقدام به تبلیغات کرده‌اند، از بخت بیشتری برای جذب سرمایه‌گذاری برخوردار بوده‌اند. در پاسخ به سؤال دیگر پژوهش در خصوص تقسیم‌بندی نوع مشخصه‌ها با توجه به اهمیت آن‌ها در پیشگویی موفقیت جذب سرمایه قبل از راه‌اندازی، مشخصه‌ها به ترتیب شامل: اول- مشخصه‌های مربوط اعتماد سرمایه‌گذاران به پیشنهاد سایت، دوم- مشخصه‌های مربوط به تجربه خلاقان و سوم- مشخصه‌های مربوط به میزان تبلیغات قبل از راه‌اندازی بود. همچنین همان‌طور که از شکل ۴ مشاهده می‌شود، عملکرد توابع مختلف پیشگویی بسیار به هم نزدیک است؛ به گونه‌ای که نمی‌توان عملکرد یکی را به طور چشم‌گیر از بقیه متمایز دانست، اما از نظر کمی قابلیت تابع خطی با ۶۲/۴٪ از بقیه بالاتر بوده است. در سناریو دوم، پیشگویی حین راه‌اندازی پروژه بوده است. به دلیل آنکه ماهیت مشخصه‌ها در این سناریو به گونه‌ای است که اطلاعات بیشتری در رابطه با موفقیت و عدم موفقیت پروژه‌ها در اختیار مدل قرار می‌دهد، لذا پیشگویی قاعده‌تاً با درصد بالاتری صورت خواهد گرفته است. در این سناریو به دلیل آنکه راه‌اندازی پروژه هم شامل مشخصه‌های حین راه‌اندازی و هم شامل مشخصه‌های قبل از راه‌اندازی است، لذا مجموع این مشخصه‌ها برای سناریو دوم در نظر گرفته شده است. در این سناریو، مدت پویا (اولین مشخصه مهم پیشگویی) بسیار مهم شناسایی شده است و نشان می‌دهد پروژه‌هایی که به دلیل اهمیت و مبلغ بالایی اولیه راه‌اندازی، زمان بیشتری را سایت برای پویا آن‌ها اختصاص داده است، شانس بالاتری نسبت به بقیه پروژه‌ها برای جذب سرمایه موردنظر خود داشته‌اند. همچنین میزان پویایی سایت شامل به‌روز شدن‌ها در خلال پویا (دومین مشخصه مهم پیشگویی حین راه‌اندازی) و تعداد کلمات پویا (چهارمین مشخصه پیشگویی در حین راه‌اندازی پروژه) عوامل مهمی هستند که نشان دهنده موفقیت جذب سرمایه پروژه خواهند بود. همچنین

مشخصه حامیان (سومین مشخصه پیشگویی حین راه‌اندازی پروژه) نوعی چشم و هم‌چشمی بین سرمایه‌گذاران برای سرمایه‌گذاری ایجاد کرده است که خود موجب موفقیت پروژه‌ها می‌شود. پس برای پاسخ به سؤال پژوهش که آیا نوع مشخصه‌های پیشگویی‌کننده بعد از راه‌اندازی با قبل از راه‌اندازی فرق می‌کند یا خیر، باید گفت بله متفاوت است و دسته‌بندی آن شامل: اول- مشخصه مربوط به مدت‌زمان پوییش، دوم- تغییر و تحول در سایت (پویایی سایت یا تعامل سازنده با سرمایه‌گذار) و سوم- چشم و هم‌چشمی برای سرمایه‌گذاری در پوییش ایجاد شده می‌شود. همان‌طور که در شکل ۵ مشاهده می‌شود، عملکرد توابع مختلف پیشگویی بسیار نسبت به هم متمایز است به‌گونه‌ای که تابع خطی با $77/37\%$ پیشگویی از بقیه بالاتر بوده است. در سناریو سوم، با دید جامع و کلی‌تری به پروژه‌های کیک‌استارتر از حیث نوع نگاه شد و این پاسخ به سؤال دیگری از سؤالات پژوهش است که آیا نوع پروژه‌ها می‌تواند به‌عنوان عامل پیشگویی در کسب موفقیت سرمایه‌گذاری مؤثر باشد یا خیر. در این رابطه، همان‌طور که در شکل ۷ دیده می‌شود، پروژه‌های مد با ۸۰ درصد، فیلم با ۷۸ درصد، موسیقی ۶۵ درصد و انتشارات با ۵۰ درصد بخت بیشتری برای جذب سرمایه دارند.

۵- پیشنهاد برای پژوهش‌های آتی

چندین مشخصه دیگر وجود دارد که تحلیل آن‌ها می‌تواند دقت پیشگویی را در حین یا قبل از راه‌اندازی پروژه‌ها، بالا ببرد. یکی از این مشخصات تحلیل متونی است که قبل یا بعد از ایجاد پوییش توسط کاربران در پاسخ به مشخصات پروژه روی سایت یا شبکه‌های اجتماعی قرار می‌گیرد. در این راستا، تحلیل و معنادار کردن این متون و ترکیب آن با مشخصات دیگر می‌تواند الگوی پیش‌بینی را با طراحی مناسب یک چارچوب هدفمند هوش کسب‌وکار موجب شود. یکی دیگر از فعالیت‌های تحقیقاتی که در این رابطه می‌توان در آینده انجام داد، تحلیل‌های زمانی برای پیشگویی روند موفقیت پروژه‌های راه‌اندازی‌شده روی سایت است. به دلیل آنکه بازه زمانی راه‌اندازی پروژه‌ها به‌صورت بازه‌ای مشخص است، می‌توان از تحلیل‌های زمانی برای عملکرد آینده سایت استفاده کرد و از آن به‌عنوان بازخورد عملکرد سایت بهره جست.

۶- منابع

- [1] Elbashir, M. Z. Collier, and , M. J. Davern, "Measuring the effects of business intelligence systems: The relationship between business process and organizational performance", International Journal of Accounting Information Systems, vol.9, 2014, pp.135-153
- [2] B. Sohrabi; I. Raeesi Vanani; F. Zareh Mirkabad., "Designing a Recommender System for Optimizing and Managing Bank Facilities through the utilization of Clustering and Classification Algorithms", Modern research in decision making ,vol. 5, 2016, pp 53-76
- [3] Van Beek, D. Den, "Performance improvement and organizational development with Business Intelligence", Tutein Nolthenius, vol.20 , 2014, pp .37-57
- [4] Juan Trujillo, Alejandro Maté , "Business intelligence 2.0: A general overview", vol. 96, 2014, pp. 98–116.
- [5] Gerber, E. M, and Hui J." Crowdfunding: Motivations and deterrents For participation," ACM Transactions on ComputerHuman Interaction, vol.20, 2016, pp.134-157
- [6] A. Xu, X. Yang, H. Rao, W. Huang, and B. P.Bailey "Show me the money!: An analysis of project updates during crowdfunding campaigns." In CHI . ACM, vol.6, 2016, pp. 591-600
- [7] Gates. S, "Aligning strategic performance measures and results" ,The Conference Board Balanced Scorecard Interest Group, vol.13, 2015, pp.100–110
- [8] L. Nazari Salari; A.Khadivar; N. Abdolvand , "A model for analyzing the barriers of using Business Intelligence (BI) in the tourism industry of Iran, a mixed method approach, Modern research in decision making" , Journal of Information Technology Management, vol.1, 2016, pp. 198-228

- [9] Alexa. <https://www.alex.com/topsites>
- [10] Dixon. B, and Candade, N, "Multispectral landuse classification using neural networks and support vectormachines: one or the other, or both?" , International J.of Remote Sensing, vol.29, 2016, pp.1185–1206.
- [11] Taeshik Shon, Yongdae Kim, Cheolwon Lee, Jongsub Moon," A machine learning framework for network anomaly detection using SVM and GA"Information Assurance Workshop, vol.6, 2013, pp.23-28