Archive of SID

Original Paper

doi: 10.22093/wwj.2020.211331.2970

Journal of Water and Wastewater, Vol. 32, No.1, pp: 69-90

An Intelligent Hybrid Model for Determining Public-Private Partnership in Iranian Water and Wastewater Industry Based on Collective Tree Algorithms

M. Eskandari[1], M. T. Taghavifard[2], I. Raeesi Vanani [3], S. S. Ghazi Noori[4]

1. PhD of Information Technology Management, Dept. of Industrial Management, College of Management and Accounting, Allameh Tabataba'I University, Tehran, Iran
(Corresponding Author) eskandary@nww.ir
2. Assoc. Prof., Dept. of Industrial Management, College of Management and Accounting, Allameh Tabataba'i University, Tehran, Iran
3. Assist. Prof., Dept. of Industrial Management, College of Management and Accounting, Allameh Tabataba'i University, Tehran, Iran
4. Assist. Prof., Dept. of Industrial Management, College of Management and Accounting, Allameh Tabataba'i University, Tehran, Iran

(Received Dec. 24, 2019 Accepted July 18, 2020)

To cite this article:
Eskandari, M., Taghavifard, M. T., Raeesi Vanani, I., Ghazi Noori, S. S. 2021. "An intelligent hybrid model for determining public-private partnership in Iranian water and wastewater industry based on collective tree algorithms" Journal of Water and Wastewater, 32(1), 69-90. Doi: 10.22093/wwj.2020.211331.2970. (In Persian)

Abstract

One of the pillars of any country's development is access to safe water and sanitation, so it is important to execute water and wastewater projects in the shortest possible time. In this regard, considering the emergence of various methods of partnership, choosing the right approach has become one of the most important issues in this industry. Therefore, a proper investment method in this field has always been the concern of decision makers. Using the database of partnership projects and data mining algorithms in the water and wastewater sector, we have designed a model to predict a proper way for public-private partnership projects. In this research, CRISP data mining method was applied to the data from 176 projects. After understanding and identifying the data, they were cleaned by deleting outliers and noisy data, and missing values were replaced. Then, the process of data classification was performed using decision tree and machine learning algorithms, and necessary analysis was performed. Also, the indicators of PPP were extracted and prioritized. K-fold cross validation method is used for validation and dividing the data. Based on the modeling and considering the data preparations and data mining methods, the stacking method is suitable for predicting and determining the type of public-private partnership contract in the implementation of each project of water and wastewater industry, which has an accuracy of 86.27%. In the pre-processing section, the combined COF method for deleting outliers and entropy factors for feature selection was used. Using the model, the success rate of each project can be predicted and an appropriate PPP contractual template for any water and wastewater project can be proposed. In addition, by entering the information of each new project, the impact of the improvement of each indicator can be easily examined.

Keywords: Water and Wastewater Industry, Public-Private Partnership, Data Mining, Forecasting, Outsourcing.

www.SID.ir

Archive of SID
dx.doi.org/10.22093/wwj.2020.211331.2970

مجله آب و فاضلاب، دوره ۳۲، شماره ۱، صفحه: ۹۰-۶۹

مدل ترکیبی هوشمند تعیین روش مشارکت عمومی- خصوصی صنعت آب و فاضلاب ایران بر مبنای الگوریتم‌های جمعی درختی

ملیحه اسکندری[۱]، محمدتقی تقوی‌فرد[۲]، ایمان رئیسی وانانی[۳]، سید سروش قاضی نوری[۴]

۱- دکترای مدیریت فناوری اطلاعات، گروه مدیریت صنعتی، دانشکده مدیریت و حسابداری، دانشگاه علامه طباطبائی، تهران، ایران
(نویسنده مسئول) eskandary@nww.ir
۲- دانشیار، گروه مدیریت صنعتی، دانشکده مدیریت و حسابداری، دانشگاه علامه طباطبائی، تهران، ایران
۳- استادیار، گروه مدیریت صنعتی، دانشکده مدیریت و حسابداری، دانشگاه علامه طباطبائی، تهران، ایران
۴- استادیار، گروه مدیریت صنعتی، دانشکده مدیریت و حسابداری، دانشگاه علامه طباطبائی، تهران، ایران

(دریافت ۹۸/۱۰/۳ پذیرش ۹۹/٤/۲۸)

برای ارجاع به این مقاله به صورت زیر اقدام بفرمایید:
اسکندری، م.، تقوی‌فرد، م. ت.، رئیسی وانانی، ا.، قاضی‌نوری، س. س.، ۱۴۰۰، " مدل ترکیبی هوشمند تعیین روش مشارکت عمومی- خصوصی صنعت آب و فاضلاب ایران بر مبنای الگوریتم‌های جمعی درختی "، مجله آب و فاضلاب، ۳۲(۱)، ۹۰-۶۹. Doi: 10.22093/wwj.2020.211331.2970

چکیده

دسترسی به آب سالم و دفع بهداشتی فاضلاب از ارکان توسعه هر کشور بوده و ضروری است که این طرح‌ها در کوتاه‌ترین زمان ممکن تکمیل شود. از سوی دیگر با توجه به ظهور انواع روش‌های مشارکت عمومی- خصوصی، انتخاب روش مناسب مشارکت، یکی از مسائل مهم بخش آب و فاضلاب بوده و این صنعت همواره نیازمند مدلی برای تصمیم‌گیری در خصوص نحوه و روش سرمایه‌گذاری در منطقه و یا طرح خاص بوده است. با توجه به وجود پایگاه داده از اطلاعات پروژه‌های مشارکتی در بخش آب و فاضلاب، می‌توان با استفاده از داده‌های به‌دست آمده و الگوریتم‌های کشف الگو و تصمیم‌گیری داده‌کاوی، مدل سرمایه‌گذاری و روش مناسب اجرای پروژه را تعیین کرد. در این پژوهش با بهره‌گیری از داده‌های ۱۷۶ پروژه مشارکتی بخش آب و فاضلاب و استفاده از روش اجرای پروژه‌های داده‌کاوی یعنی فرایند کریسپ، مدل مشارکت عمومی- خصوصی صنعت آب و فاضلاب استخراج شد. پس از تشریح و درک داده، مراحل پاک‌سازی و حذف داده‌های پرت اجرا شده است. در مرحله دسته‌بندی، با تکنیک‌های درختی و یادگیری ماشین، طبقه‌بندی موفقیت و شکست پروژه‌ها و تحلیل‌های لازم انجام و شاخص‌های مشارکت عمومی- خصوصی به‌ترتیب اولویت استخراج شد. به‌منظور اعتبارسنجی و تقسیم داده‌ها، از روش اعتبارسنجی ضربدری استفاده شد. بر مبنای مدل‌سازی انجام شده و با در نظر گرفتن انواع روش‌های پیش‌پردازش و داده‌کاوی، روش استکینگ با دقت ۸۶/۲۷ درصد، به‌عنوان روش مناسب پیش‌بینی و تعیین نوع قرارداد مشارکت عمومی- خصوصی در اجرای هر پروژه بخش آب و فاضلاب معرفی شد. در بخش پیش‌پردازش نیز روش ترکیبی Connectivity-based Outlier Factor برای حذف داده پرت و شاخص آنتروپی برای انتخاب ویژگی استفاده شد. با توجه به مدل پیشنهادی، علاوه بر معرفی قالب قراردادی مشارکتی مناسب برای اجرای هر گروه از پروژه‌های بخش آب و فاضلاب و معرفی قالب قراردادی مناسب در هر استان، می‌توان میزان موفقیت هر طرح را در هر یک از قالب‌های قراردادی پیش‌بینی و تأثیر بهبود هر یک از شاخص‌ها را در افزایش موفقیت پروژه بررسی کرد.

واژه‌های کلیدی: صنعت آب و فاضلاب، مشارکت عمومی- خصوصی، داده‌کاوی، پیش‌بینی، برون‌سپاری

www.SID.ir

dx.doi.org/10.22092/ajwwj.2020.211331.2970
Archive of SID

۱- مقدمه

در دنیای امروز و با توجه به پیشرفت‌های روزافزون اطلاعات محور، لزوم استفاده از فناوری اطلاعات و ارتباطات در برنامه‌ریزی و تصمیم‌گیری بر کسی پوشیده نیست. صنعت آب و فاضلاب نیز از این قاعده مستثنا نبوده و همواره سعی داشته تا با به‌کارگیری مدل‌ها و ابزارهای جدید، کارایی را در حوزه‌های کاری مختلف خود بهبود بخشد (Mutula and Van Brakel, 2006).

از جمله مسائلی که موجب اختلال در تکمیل طرح‌های این صنعت می‌شود، کمبود منابع مالی است. بر مبنای مستندات بانک جهانی[1]، سازمان بهداشت جهانی[2] و یونیسف[3]، برای بهره‌مندی مردم از سیستم دفع بهداشتی فاضلاب در کشورهای آسیایی در حدود ۳۰۰ میلیارد دلار اعتبار نیاز است (Hall et al., 2005).

با توجه به وابستگی این طرح‌ها به منابع دولتی در اکثر این کشورها، تحقق این امر با مشکلات فراوانی روبه‌رو است. در واقع، حجم زیاد منابع مالی موردنیاز برای رسیدن به اهداف برنامه‌ریزی شده توسعه‌ای و محدودیت منابع مالی دولتی در سالیان اخیر، رویکرد استفاده از روش‌های نوین تأمین منابع مالی و سرمایه‌گذاری بخش غیردولتی برای تأمین منابع مالی موردنیاز طرح‌های بخش آب و فاضلاب را نه‌تنها یک راه‌حل، بلکه تبدیل به یک ضرورت کرده تا بتوان از توان اقتصادی، مدیریتی و اجرایی بخش خصوصی کشور استفاده کرد.

شرکت‌های آب و فاضلاب بر مبنای رسالت اصلی خود، مسئول تأمین آب سالم و دفع بهداشتی فاضلاب هستند که متأسفانه در این مسیر و به‌منظور تأمین منابع مالی، همواره با مشکلاتی مواجه بوده‌اند. از یک سو هزینه‌بری زیاد طرح‌های آب و فاضلاب که بر مبنای پژوهش‌های انجام شده در این صنعت به میزان ۲۰۶ هزار میلیارد ریال در بخش شهری و ۲۵ هزار میلیارد ریال در بخش روستایی، در کل جهان است (Company, 2009).

از سوی دیگر محدودیت‌های منابع مالی، اجرایی و تخصصی در شرکت‌های آب و فاضلاب و سایر محدودیت‌ها، باعث شده که اجرای پروژه‌ها توسط بخش دولتی معمولاً در زمان طولانی انجام شده و هزینه‌های اضافی بیشتری نیز به پروژه تحمیل شود. بر این اساس، ورود بخش خصوصی به طرح‌های صنعت آب و فاضلاب در قالب سرمایه‌گذار، مجری و بهره‌بردار واجب است. در کشورهای توسعه‌یافته از دهه ۱۹۸۰، به‌طور فزاینده‌ای برای احداث و توسعه تأسیسات زیربنایی به راهکارهای مشارکتی روی آورده‌اند که این امر در اوایل اجرای آن به‌دلیل تأمین بودجه و تمایل برای بازدهی بیشتر طرح‌ها بوده است.

با توجه به وجود انواع روش‌های مشارکت بخش خصوصی، انتخاب روش مناسب که با مقتضیات پروژه، محل اجرای آن و سایر عوامل مرتبط با اجرای طرح سازگاری داشته باشد، به یکی از مسائل مهم در این صنعت تبدیل شده و همواره نیازمند مدلی است تا به‌راحتی به مدیران ارشد و سرمایه‌گذاران در انتخاب روش مناسب برای هر طرح و یا منطقه خاص کمک کند. به‌منظور تحقق این امر و ارائه مدل مشارکت عمومی- خصوصی در صنعت آب و فاضلاب، ضروری است تا پس از شناسایی نیازهای این صنعت، انواع روش‌های مشارکت بخش خصوصی در اجرای طرح‌ها شناسایی شوند.

در قدم بعدی ضروری است تا عواملی که در جذب سرمایه‌گذاران تأثیرگذار هستند، بررسی شد و ارتباط میان آنها مشخص شود. در نهایت و با استفاده از اطلاعات به‌دست‌آمده و الگوریتم‌های داده‌کاوی در حوزه انتخاب ویژگی و طبقه‌بندی با روش‌های جنگل تصادفی، گرادیان بوستینگ[4] و استکینگ می‌توان مدل سرمایه‌گذاری بخش خصوصی در این صنعت را ترسیم کرده و در خصوص هر یک از طرح‌ها و مناطق روش مناسب اجرای پروژه را پیش‌بینی کرد. ازاین‌رو در ادامه ابتدا مروری بر کارهای انجام شده و اهداف مشارکت بخش خصوصی و روش‌های داده‌کاوی، خواهیم داشت. سپس تعاریف و الگوریتم‌های داده‌کاوی کار شده در بخش دسته‌بندی پروژه‌ها را تشریح کرده و بر اساس روش پژوهش، مدل پیشنهادی در بخش پیش‌پردازش و پاک‌سازی، بررسی شاخص‌های تأثیرگذار با سود اطلاعاتی و مرحله دسته‌بندی ارائه خواهد شد و در نهایت نتایج اجرا، گزارش و تحلیل می‌شود.

۲- پیشینه پژوهش

مشارکت عمومی- خصوصی[5] رویکردی است که برای افزایش

[1] World Bank
[2] World Health Organization (WHO)
[3] UNICEF

[4] Boosting
[5] Public-Private Partnership (PPP)

www.SID.ir

Archive of SID
dx.doi.org/ 10.22093/wwj.2020.211331.2970

ارزش اقتصادی خروجی زیرساخت‌ها استفاده می‌شود و شامل طیف وسیعی از زیرساخت‌های بخش عمومی می‌شود (Cui et al., 2018).

این مشارکت یک قرارداد بلندمدت بین یک سازمان عمومی و نماینده خصوصی است که به موجب آن منابع و ریسک‌های پروژه بین طرفین قرارداد تقسیم می‌شود و هدف از این قراردادها، توسعه و یا نوسازی تسهیلات عمومی است (Mcquaid and Scherrer, 2010).

در سال‌های اخیر به‌دلیل رشد روزافزون جمعیت و توسعه اقتصادی کشورهای مختلف، نیاز شدیدی به ساختارهای زیربنایی در بسیاری از کشورها احساس می‌شود (Hodge et al., 2010) و این در حالی است که بودجه‌های دولتی برای توسعه زیرساخت‌های موردنیاز یک کشور به‌صورت متعارف محدود و غیربهینه است. به همین دلیل دولت‌ها همواره تلاش می‌کنند با به‌کارگیری بخش خصوصی و فعال کردن این بخش در پروژه‌های زیربنایی، راه‌حلی نوین برای این کاستی‌ها بیابند.

در این پژوهش، ضروری است که ابتدا اهداف مشارکت بخش خصوصی در صنعت آب و فاضلاب و در نتیجه شاخص‌های مؤثر در ورود بخش خصوصی به این طرح‌ها شناسایی شوند. علاوه بر این باید اقدامات انجام شده در زمینه استفاده از داده‌کاوی در مدل‌سازی و پیش‌بینی بررسی شود. با توجه به‌این امر، می‌توان پیشینه پژوهش را در دو بخش مطرح کرد:

- شناسایی اهداف و در نتیجه استخراج شاخص‌های مشارکت بخش خصوصی در صنعت آب و فاضلاب.
- شناسایی و بررسی پژوهش‌های انجام شده در زمینه استفاده از داده‌کاوی در پیش‌بینی.

۳- شناسایی اهداف و شاخص‌های مشارکت عمومی- خصوصی در صنعت آب و فاضلاب ایران

در سال‌های اولیه و با توجه به نوظهور بودن روش‌های مشارکت عمومی- خصوصی، مهم‌ترین شاخص‌های این‌گونه طرح‌ها را تکمیل طرح در زمان مقرر و ایجاد حداکثر نتایج عنوان می‌کردند (Yu et al., 2007).

در سال‌های بعد و با توجه به تجربیات به‌وجود آمده در اجرای طرح‌های مختلف به‌ویژه در حوزه ایجاد زیرساخت‌های شهری، انتخاب روش بهینه به‌منظور موفقیت اجرای پروژه‌ها، مورد توجه پژوهشگران قرار گرفت و شاخص‌های متعددی به‌منظور بررسی اجرای طرح‌ها به روش‌های مختلف موردنظر گرفته شد. با توجه به رشد بالای استفاده از روش‌های مشارکت عمومی- خصوصی از سال ۲۰۰۰ به بعد، در این پژوهش، مبنای پژوهش‌ها در شناسایی شاخص‌های مشارکت عمومی- خصوصی، سال ۲۰۰۰ در نظر گرفته شد و شاخص‌های مربوطه بر مبنای این پژوهش‌ها استخراج شد.

از جمله پژوهش‌هایی که در زمینه شناسایی شاخص‌های مشارکت انجام شده است می‌توان به پژوهش‌های کیائو و همکاران در سال ۲۰۰۱ اشاره کرد که شاخص‌های حیاتی موفقیت[1] در پروژه‌های BOT[2] بررسی شد و ۶ شاخص اصلی نوآوری و کارآفرینی، انتخاب پروژه مناسب، تشکیل ترکیبی قوی از سهام‌داران، راهکارهای تکنیکی مناسب، وجود پروپوزال مالی قوی و شرایط ویژه مناقصه را به‌عنوان عوامل حیاتی موفقیت عنوان کردند (Qiao et al., 2001).

در سال ۲۰۰۳ آکینتوی و همکاران، مهم‌ترین شاخص‌های مشارکت بهینه بخش عمومی- خصوصی در اجرای پروژه‌ها را بهبود کیفیت، افزایش کارایی و اثربخشی، استفاده بهینه از منابع مالی و بهره‌مندی از استانداردهای بهتر کارایی، عنوان کردند (Akintoye et al., 2003).

در پژوهشی دیگر، ژنگ و همکاران پس از بررسی و تجمیع کلیه بررسی‌هایی که تا این سال بر روی شناسایی شاخص‌های مشارکت عمومی- خصوصی انجام شده بود، CSFهای اصلی مشارکت عمومی- خصوصی در پروژه‌های زیرساختی را به‌صورت وجود محیط مساعد برای سرمایه‌گذاری، قابلیت حیات اقتصادی[3]، تسهیم ریسک مناسب و وجود قراردادهای مطمئن، وجود مشارکت سرمایه‌گذاران بانفوذ و دارای قدرت تکنیکی زیاد و ارائه بسته‌های مالی مناسب، ارائه کردند (Zhang, 2005).

لی و همکاران نیز با بررسی پروژه‌های مشارکتی انجام شده در سال ۲۰۰۵ ، CSFهای مهم در اجرای طرح‌های مشارکت عمومی- خصوصی را وجود ترکیب مناسب و قوی در بخش خصوصی،

[1] Critical Success Factors (CSF)
[2] Build- Operate- Transfer (BOT)
[3] Economic Viability

www.SID.ir

dx.doi.org/10.22092/wwj.2020.211331.2970
Archive of SID

ارزیابی واقع‌بینانه هزینه - فایده، وجود بازارهای مالی در دسترس، تقسیم و تسهیم مناسب ریسک، امکان‌سنجی تکنیکی پروژه، حمایت سیاسی، وجود فرایند رقابتی در تدارکات، وجود شفافیت در فرایند تدارکات، وجود اهداف و فواید چندبعدی، تعهد و مسئولیت‌پذیری بخش عمومی و خصوصی، حکمرانی مناسب، مشارکت دولت در تأمین ضمانت‌های لازم، وجود اقتصاد کلان باثبات، وجود چارچوب مناسب قانونی، شرایط سیاسی و اقتصادی، وجود ساختارهای خصوصی مناسب، تقسیم حکمرانی بین بخش عمومی و خصوصی، پشتیبانی اجتماعی و انتقال تکنولوژی، عنوان کردند (Li et al., 2005).

بوهل و همکاران با انجام پژوهشی در زمینه مشارکت عمومی- خصوصی در صنعت آب و فاضلاب، اقدام به شناسایی این شاخص‌ها در کشور آلمان کردند. این گروه با شناسایی اهداف مشارکت عمومی - خصوصی در این صنعت، ۳۶ شاخص را به‌عنوان شاخص‌های کلیدی کارایی و موفقیت پروژه‌های PPP اعلام کرد و از طریق بررسی پروژه‌های انجام شده در برخی از کشورها به وزن‌دهی شاخص‌ها پرداختند (Böhl, 2007).

ساش و همکاران نیز طی پژوهشی به بررسی فرصت‌ها و تأثیرات ریسک‌های سیاسی بر روی پروژه‌های مشارکت عمومی- خصوصی حوزه زیرساخت‌ها در چین و ۱۴ کشور دیگر آسیا پرداختند (Sachs et al., 2007). بسیاری از این پروژه‌ها در حوزه آب و فاضلاب و به‌ویژه ساخت تصفیه‌خانه‌های آب و فاضلاب با مشارکت عمومی- خصوصی بود و در طی آن تأثیر ۶ فاکتور و ریسک سیاسی عدم‌سهولت تسعیر ارز و محدودیت‌های تبادل آن، سلب مالکیت‌ها و انحصارهای دولتی، عدم‌ایفای تعهدات، تهدیدها و مشکلات سیاسی، ریسک‌های حقوقی و قانونی ناشی از بوروکراسی و ریسک‌های غیرمتأثر از دولت، بر روی این گروه از طرح‌ها بررسی شد. در این پژوهش، وضعیت هر یک از کشورها در ۶ حوزه مورد اشاره، بررسی شد و راهکارهای مناسب در هر یک از کشورها را مطرح کرد.

در همان سال، سالمن و همکاران بررسی‌هایی را در پروژه‌های مشارکت عمومی- خصوصی و به‌ویژه پروژه‌های BOT حوزه‌ی زیرساخت انجام دادند و مهم‌ترین شاخص‌های مؤثر در تصمیم‌گیری و موفقیت را در بخش‌های مختلف استخراج و دسته‌بندی کردند. این گروه ۲۱ شاخص را در ۳ دسته عوامل محیطی و حقوقی، عوامل فنی و تکنیکی و عوامل مالی و تجاری مطرح کردند (Salman et al., 2007).

یان و همکاران انتخاب مدل مشارکت عمومی- خصوص را از دید کلیه ذی‌نفعان پروژه بررسی کردند. ذی‌نفعان مورد بررسی در این پژوهش ۴ بخش اصلی جامعه، بخش خصوصی، بخش عمومی و دولتی و جامعه دانشگاهی را شامل می‌شد. این پژوهش ۱۵ شاخص را به‌عنوان اصلی‌ترین شاخص‌های عملکردی در این پروژه‌ها مطرح کرد و اهمیت این شاخص‌ها را در هر یک از گروه‌های ذی‌نفع به تفکیک مشخص کرد. وی در ادامه با استفاده از روش‌های فازی شاخص‌های مورد اشاره را وزن‌دهی و اهمیت آنها را در هر یک از گروه‌های ذی‌نفع مشخص کرد (Yuan et al., 2009).

در پژوهشی دیگر با بررسی پروژه‌هایی که در حوزه مشارکت عمومی- خصوصی انجام شد، شاخص‌های کلیدی موفقیت را در پنج حوزه اصلی تقسیم کردند و با استفاده از پرسش‌نامه و ارائه آن به ذی‌نفعان درگیر در پروژه‌های مشارکت عمومی- خصوصی، در مجموع ۴۸ شاخص را به‌عنوان عوامل کلیدی کارایی در پروژه‌های مشارکتی شناسایی کرد و پس از شناسایی، به‌منظور تعیین اهمیت هر یک از شاخص‌ها نیز از پرسش‌نامه‌های مبتنی بر طیف لیکرت[1] استفاده شد (Yuan et al., 2011).

گلابچی پژوهشی را با عنوان انتخاب روش مناسب مشارکت خصوصی- دولتی به‌منظور تأمین مالی پروژه‌های آزادراه ایران انجام داد و پس از ارائه دو پرسش‌نامه به خبرگان حوزه مورد پژوهش، تحلیل نتایج با استفاده از روش آماره تی، ۱۹ معیار مهم در انتخاب روش مناسب مشارکت عمومی- خصوصی را عنوان کردند (Golabchi, 2015).

بانک جهانی نیز در گزارش خود با عنوان ایجاد چارچوب برای مشارکت عمومی- خصوصی در سال ۲۰۱۵، اصلی‌ترین محدودیت‌ها در پیشبرد طرح‌های مشارکتی را محدودیت‌های قانونی عنوان کرد و عوامل مهم را در قالب ۱۵ مورد بیان کرده است (Delmon, 2015).

مرزوک و فائز نیز از طریق استفاده از پرسش‌نامه در حوزه مشارکت عمومی- خصوصی، ۴۶ شاخص را در ۶ بخش مطرح کردند (Marzouk and Fayez, 2018).

[1] Likert-Style Rating Questions

www.SID.ir

Archive of SID
dx.doi.org/ 10.22093/wwj.2020.211331.2970

اگرچه در سال‌های اخیر پژوهش‌های بسیار زیادی بر روی قراردادهای PPP انجام‌شد، اما بررسی‌های انجام شده بر روی این پژوهش‌ها نشان داد شناسایی و ارزیابی معیارهای مؤثر به‌منظور انتخاب نوع قرارداد استفاده شده برای تأمین مالی پروژه‌های صنعت آب و فاضلاب در دنیا و به‌ویژه در ایران، کمتر مورد توجه قرار گرفته است.

بر اساس پژوهش‌های انجام شده در صنعت آب و فاضلاب در سایر کشورها و تطابق آن با پژوهش‌های انجام شده در شرکت مهندسی آب و فاضلاب کشور، مشارکت بخش خصوصی در طرح‌های آب و فاضلاب بر مبنای ۷ هدف کلی دانش و تجربیات تکنیکی، اثربخشی عملیاتی، پاسخ‌گویی به مشتریان، کارایی اقتصادی، سرمایه‌گذاری، دانش و تجربیات مدیریتی و مصون ماندن از مداخلات سیاسی انجام می‌شد (Böhl, 2007, Company, 2009).

به‌منظور استفاده از این اهداف و شناسایی شاخص‌های مشارکت عمومی- خصوصی، می‌توان برای سنجش میزان تحقق هر یک از اهداف، تعدادی از شاخص‌های کلیدی عملکرد[1] را قرار داد و با توجه به نتیجه پژوهش‌های انجام شده در سایر کشورها و در قالب هر یک از اهداف هفت‌گانه مشارکت بخش خصوصی، می‌توان کلیه KPIهای مؤثر و موجود را شناسایی کرد (Alegre and Association, 2006).

پژوهش‌های انجام شده در زمینه استفاده از داده‌کاوی در پیش‌بینی داده‌کاوی در حوزه‌های زیادی ازجمله بازاریابی، تأمین منابع مالی، بانکداری، تولید، سلامت، مدیریت روابط مشتریان، تشخیص خطا، پیش‌بینی و یادگیری سازمانی استفاده شد که در ادامه به اختصار به برخی از این پژوهش‌ها اشاره شده است.

بیک‌زاده و همکاران از داده‌کاوی در استخراج قواعد و طبقه‌بندی داده‌ها به‌منظور پیش‌بینی و کمک به بهبود فرایند تصمیم‌گیری در دستگاه‌های آموزشی و مؤسسات آموزش عالی پرداختند (Beikzadeh et al., 2008).

الردیده و همکاران از داده‌کاوی و تکنیک‌های طبقه‌بندی، برای بهبود عملکرد دستگاه‌های آموزشی در مؤسسات آموزش عالی استفاده کردند. در این پژوهش در ابتدا عوامل اصلی اثرگذار بر کارایی و عملکرد دانشجویان شناسایی و اطلاعات آنها جمع‌آوری شد و در نهایت با استفاده از درخت تصمیم و داده‌های مربوط به هر دانشجو، پیش‌بینی عملکرد و نمرات دانشجو انجام شد (Al-Radaideh and Al Nagi, 2012).

این مطلب توسط کریمی و موتوری نیز ارزیابی شد که در این پژوهش بسیاری از داده‌های کیفی نیز به مطالعه افزوده شد (Kirimi and Moturi, 2016).

لاوراک و همکاران از الگوریتم‌های طبقه‌بندی در خصوص پیش‌بینی و بررسی لزوم مراجعه بیماران به متخصصین استفاده کردند. در این پژوهش، دسته‌بندی بیماران در مراکز بهداشت از خدمات اولیه سلامت تا رجوع به متخصصین، که ترکیبی از داده‌کاوی و رویکردهای پشتیبان تصمیم برای تشخیص لزوم مراجعه بیمار به متخصص است، بررسی شد و در نهایت یک واسط کاربری مناسب برای کمک به مدیریت دانش و پشتیبانی تصمیم ارائه شد (Lavrač et al., 2007).

چنگ و همکاران از الگوریتم‌های طبقه‌بندی، در طراحی سیستم مدیریت دانش مالی[2] استفاده کردند و از ترکیب الگوریتم‌های خوشه‌بندی و تکنیک‌های طبقه‌بندی داده‌کاوی برای طبقه‌بندی داده‌های گروهی و منسجم استفاده کردند (Cheng et al., 2009).

چن و دو نیز در همان سال از تکنیک‌های داده‌کاوی و شبکه‌های عصبی به‌منظور پیش‌بینی بحران‌های مالی بهره بردند. در این پژوهش، مدلی طراحی شد که با دریافت اطلاعات مربوط به شرکت‌ها، وضعیت مالی آنها را بررسی کردند و احتمال بروز بحران مالی را پیش‌بینی کردند (Chen and Du, 2009).

لی و همکاران نمونه‌ای از الگوریتم طبقه‌بندی را در صنایع کوچک و متوسط پیاده‌سازی کردند. در این روش، ترکیبی از داده‌کاوی و مدیریت دانش به‌منظور طراحی یک سیستم پشتیبان تصمیم استفاده شد. در این پژوهش که به‌منظور پیش‌بینی و بهینه‌سازی زمان رسیدن غذا در زنجیره تأمین یک شرکت تأمین غذا انجام شد، ابتدا عوامل مؤثر بر روی زمان رسیدن غذا شناسایی و سپس با استفاده از درخت تصمیم و شبکه‌های عصبی، زمان رسیدن غذا پیش‌بینی شد (Li et al., 2010).

ریبیرو و همکاران از تکنیک‌های داده‌کاوی به‌منظور بررسی و

[1] Key Performance Indicators (KPI)

[2] Financial Knowledge Management System (FKMS)

www.SID.ir

Archive of SID
DOI: 10.22093/wwj.2020.211331.2970

پیش‌بینی کارایی در تصفیه‌خانه‌های فاضلاب استفاده کردند. در این پژوهش از تکنیک‌های SVM[1] بر روی داده‌های موجود استفاده شد و متغیرهای موردنیاز و در نتیجه کارایی تصفیه‌خانه پیش‌بینی شد (Ribeiro et al., 2013).

جنابی و میروشاندل از این تکنیک برای بهبود مدیریت امور مشتریان استفاده کردند (Jenabi and Mirroshandel, 2014).

سان و همکاران نیز میزان شکستگی و اتفاقات انجام شده در شبکه‌های آب را توسط خوشه‌بندی فضایی تجزیه‌وتحلیل کردند و سپس با خوشه‌بندی مبتنی بر چگالی، محدوده‌هایی از شبکه تأمین آب که در معرض اتفاقات هستند را شناسایی کردند (Sun et al., 2014).

کیو و همکاران در پژوهشی، یک سیستم مبتنی بر داده‌کاوی را طراحی و پیاده‌سازی کردند که فرایند طراحی و ساخت تصفیه‌خانه‌های فاضلاب را در کشور چین بهینه کند (Qiu et al., 2018).

علاوه بر آنچه در بالا بیان شده است، در پژوهش‌های بسیاری نیز پیش‌بینی کیفیت آب با استفاده از الگوریتم‌های داده‌کاوی انجام شده است (Azhar et al., 2013, Ogwueleka and Ogwueleka, 2010, Ali and Qamar, 2013, Wen et al., 2013).

بر مبنای بررسی‌های انجام شده، مشاهده می‌شود که شناسایی شاخص‌های مشارکت عمومی- خصوصی در صنعت آب و فاضلاب صرفاً در پژوهش بوهل انجام شده که با مقتضیات کشور ایران اختلافات بسیاری دارد ولی متدلوژی آن در کنار شاخص‌های استخراج شده از سایر پژوهش‌ها، قابل‌استفاده در صنعت آب و فاضلاب ایران است. در زمینه استفاده از داده‌کاوی در حوزه مشارکت عمومی- خصوصی و مدل‌سازی آن، تاکنون مورد مشابهی در ایران و جهان یافت نشده است ولی بررسی تجربیات در سایر حوزه‌ها می‌تواند به انتخاب روش و الگوریتم مناسب کمک کند (Böhl, 2007).

۳-۱- مبانی نظری

در این پژوهش از تکنیک‌های داده‌کاوی برای پژوهش دو هدف اصلی استفاده می‌شود. در ابتدا لازم است که بتوان موفقیت و یا عدم موفقیت یک پروژه را در قالب قراردادی مشخص پیش‌بینی کرد و در مرحله بعد بتوان عدد مربوط به میزان موفقیت را به‌عنوان خروجی ارائه کرد. در این پژوهش الگوریتم‌های مختلفی استفاده شده‌اند که شامل گرادیان بوستینگ، شبکه عصبی، رگرسیون لاجستیک[2]، یادگیری عمیق[3] و بگینگ[4] و استکینگ[5] است.

۳-۲- شبکه عصبی

شبکه‌های عصبی مصنوعی الگویی برای پردازش اطلاعات هستند که با تقلید از شبکه‌های عصبی بیولوژیکی مثل مغز انسان ساخته شده‌اند. عنصر کلیدی این الگو ساختار جدید سیستم پردازش اطلاعات آن است و از تعداد زیادی عناصر (نرون) با ارتباطات قوی داخلی که هماهنگ باهم برای حل مسائل مخصوص کار می‌کنند تشکیل شده است. شبکه‌های عصبی مصنوعی با پردازش روی داده‌های تجربی، دانش یا قانون نهفته در ورای داده‌ها را به ساختار شبکه منتقل می‌کند که به این عمل یادگیری می‌گویند. هر شبکه عصبی شامل یک لایه ورودی است که هر گره در این لایه معادل یکی از متغیرهای پیش‌بینی است. گره‌های موجود در لایه میانی به تعدادی گره در لایه نهان، وصل می‌شوند. هر گره ورودی به همه گره‌های لایه نهان متصل می‌شود. برای شبیه‌سازی یک روش طبقه‌بندی بر روی شبکه‌های عصبی در لایه ورودی، باید به تعداد صفات خاصه نرون قرار دهیم. در لایه خروجی نیز باید به تعداد کلاس‌ها نرون داشته باشیم.

۳-۳- درخت تصمیم

درخت تصمیم در داده‌کاوی مدلی است که برای نمایش رده‌بندها و رگرسیون استفاده می‌شود. درختان تصمیم موضوع‌ها را بر اساس مقدار صفت‌ها رده‌بندی می‌کنند. برگ‌ها نماد پیش‌بینی هستند، هر گره در یک درخت تصمیم نماینده صفت در یک موضوع مورد رده‌بندی است و هر شاخه نماینده مقداری است که یک گره می‌تواند اختیار کند. این رده‌بندهای مبتنی بر قوانین اگر- آنگاه، یکی از فن‌های ساده و قابل‌فهمی است که در موضوعات مختلف استفاده می‌شود.

[1] Support Vector Machines (SVM)

[2] Logistic Regression
[3] Deep Learning
[4] Bagging
[5] Stacking

www.SID.ir

dx.doi.org/ 10.22093/wwj.2020.211331.2970
Archive of SID

از جمله مفاهیم مطرح در این حوزه، آنتروپی است که میزان خلوص (بی‌نظمی یا عدم‌خالص بودن) مجموعه‌ای از مثال‌ها را مشخص می‌کند و هرچه این مقدار بیشتر باشد، یعنی داده‌ها اطلاعات کمتری را در خود دارند. اگر مجموعه S شامل مثال‌های مثبت و منفی از یک مفهوم هدف باشد، آنتروپی S نسبت به این دسته‌بندی بولی به‌صورت معادله ۱ تعریف می‌شود

$$\log_2^{p^-} \times P^- - \log_2^{p^+} \times \text{Entropy}(S) = -P^+ \qquad (۱)$$

۳-۴- آدابوست[1]

از روش‌های جمعی یا ترکیبی است که یک مجموعه از دسته‌بندی‌های ضعیف را گرفته و خروجی آنها را با یکدیگر ترکیب می‌کنند تا دسته‌بندی نهایی را به‌گونه‌ای بسازد که کارایی آن از کارایی تک‌تک دسته‌بندی‌های استفاده‌شده در الگوریتم، بیشتر باشد. در نهایت دسته رکوردهای دیده نشده در مرحله ارزیابی را با ترکیب کردن خروجی تک‌تک دسته‌بندی‌های کوچک استفاده‌شده، تعیین می‌کند. آدابوست مخفف بوستینگ تطبیقی است و یک الگوریتم یادگیری ماشین است که توسط یاو فروند و رابرت شاپیر ابداع شد (Freund and Schapire, 1995). در واقع آدابوست یک متا الگوریتم است که به‌منظور ارتقای عملکرد و رفع مشکل رده‌های نامتوازن[2]، همراه دیگر الگوریتم‌های یادگیری استفاده می‌شود. در این الگوریتم، طبقه‌بندی در هر مرحله جدید به‌نفع نمونه‌های غلط طبقه‌بندی شده در مراحل قبل تنظیم می‌شود. آدابوست نسبت به داده‌های نویزی و پرت، حساس است.

۳-۵- پشته‌سازی یا استکینگ

مجموعه‌ای از طبقه‌بندی‌کننده‌ها که در آن با استفاده از نمونه بوت استراپ داده‌های آموزشی، آموزش داده می‌شوند. ایده اصلی این است که ببینیم آیا داده‌های آموزشی به‌خوبی آموزش دیده است یا نه. مثلاً اگر یک طبقه‌بندی‌کننده خاص در منطقه خاصی از ویژگی‌ها اشتباه آموزش داده شده باشد و به‌همین دلیل نمونه‌های آن منطقه را به اشتباه دسته‌بندی کرده باشد، ردیف ۲ طبقه‌بندی‌کننده‌ها قادر به یادگیری این رفتار بوده و می‌تواند با استفاده از این رفتاری که یاد گرفته و رفتارهای یاد گرفته شده از دیگر طبقه‌بندی‌کننده‌ها، چنین آموزش نادرستی را اصلاح کند و ارزیابی کراس نیز برای آموزش طبقه‌بندی‌کننده‌های ردیف اول استفاده می‌شود. کل داده‌های آموزشی به T بلوک تقسیم می‌شوند و هر طبقه‌بندی‌کننده ردیف اول، اولین بار بر روی مجموعه‌ای متفاوت از T-1 بلوک از داده‌های آموزشی آموزش داده می‌شود. سپس هر طبقه‌بندی‌کننده بر روی T اُمین بلوک که در طول آموزش با آن مواجه نشده است، ارزیابی می‌شود. خروجی این طبقه‌بندی‌کننده‌ها همراه با برچسب واقعی و درست برای آن بلوک‌ها، مجموعه داده آموزشی برای ردیف ۲ طبقه‌بندی‌کننده‌ها را تشکیل می‌دهند (Rokach, 2010).

۳-۶- جنگل تصادفی[3]

جنگل تصادفی درخت تصمیم‌های زیادی را تولید می‌کند. برای طبقه‌بندی یک شیء جدید، بردار ورودی را در انتهای هر یک از درختان جنگل تصادفی قرار می‌دهد. هر درخت یک طبقه‌بندی را ارائه می‌کند و اصطلاحاً گفته می‌شود این درخت به آن کلاس «رأی» می‌دهد. جنگل، طبقه‌بندی که بیشترین رأی را داشته باشد (بین همه درخت‌های جنگل) انتخاب می‌کند (Kirimi and Moturi, 2016). هر درخت به‌صورت زیر تشکیل می‌شود:

۱- اگر N تعداد حالت‌ها در مجموعه داده‌های آموزش باشد، N حالت را به‌صورت تصادفی با جای‌گذاری از داده‌های اصلی، نمونه‌گیری می‌کنیم. این نمونه، مجموعه کار برای این درخت است.

۲- اگر M متغیر داشته باشیم و m را کوچکتر از M در نظر بگیریم به‌طوری که در هر گره، m متغیر به‌صورت تصادفی از M انتخاب می‌شوند و بهترین جداسازی روی این m متغیر برای جداسازی گره استفاده می‌شود. مقدار m در طول ساخت جنگل ثابت در نظر گرفته می‌شود. هر درخت به اندازه ممکن بزرگ می‌شود. هیچ هرسی وجود ندارد.

۳-۷- گرادیان بوستینگ

روش گرادیان بوستینگ مانند جنگل تصادفی از درخت‌های تصمیم «ضعیف» استفاده می‌کند. تفاوت بزرگ این دو روش آن است که در روش گرادیان بوستینگ، درخت‌ها یکی پس از دیگری آموزش داده می‌شوند. هر درخت زیرمجموعه، در درجه اول با داده‌هایی که

[1] Adaboost
[2] Class Imbalance
[3] Random Forest

www.SID.ir

Archive of SID

به اشتباه توسط درخت پیشین پیش‌بینی شده‌اند آموزش داده می‌شوند. این امر موجب می‌شود مدل کمتر بر مسائلی که پیش‌بینی در آنها آسان است تمرکز داشته و بیشتر روی موارد پیچیده متمرکز شود.

۳-۸- روش پژوهش

در این پژوهش به‌منظور شناسایی روش‌های مشارکت بخش خصوصی در طرح‌های آب و فاضلاب و نیز عوامل مرتبط با آنها، از مطالعات کتابخانه‌ای و روش‌های میدانی بهره برده شده و در ادامه به‌منظور درک تأثیرات متقابل هر یک از این عوامل بر روی یکدیگر و در نهایت ارائه مدل مفهومی مشارکت بخش خصوصی در طرح‌های آب و فاضلاب، از الگوریتم‌های داده‌کاوی استفاده شد. به‌منظور انجام داده‌کاوی از متدلوژی CRISP-DM استفاده شد و پیاده‌سازی با نرم‌افزار Rapid miner نسخه ۹ انجام شد. در ادامه روش پژوهش بر مبنای متدلوژی کریسپ به اختصار تشریح می‌شود:

شناخت بازار (درک فضای کسب و کار): بر مبنای اطلاعات موجود در وب‌سایت رسمی تجهیز منابع مالی و توسعه مشارکت بخش غیردولتی صنعت آب و فاضلاب[1]، این بخش به‌عنوان صنعت پیشرو در حوزه ارائه روش‌های نوین جذب مشارکت بخش خصوصی، همواره در سطح کشور مطرح بوده است و توانسته با ابداع قالب‌های مختلف قراردادهای مشارکت عمومی- خصوصی و انعقاد قراردادهایی در هر قالب، تعداد بسیاری از پروژه‌های زیرساختی این حوزه را در کمترین زمان و با بیشترین کیفیت به بهره‌برداری رساند. با توجه به گذشت بیش از ۱۷ سال از انعقاد قراردادهای مشارکتی در این صنعت، تجربه، آمار و اطلاعات مناسبی در این حوزه وجود داشته که می‌تواند کمک شایانی به پیشبرد پژوهش کند. در حال حاضر، دغدغه فعلی تصمیم‌گیران در این حوزه، انتخاب روش مناسب مشارکت عمومی- خصوصی در پیشبرد طرح‌ها است که با توجه به حجم زیاد سرمایه‌گذاری مورد نیاز و اهمیت بسیار زیاد اجرای این طرح‌ها در زندگی افراد جامعه، این موضوع به یکی از دغدغه‌های اصلی در این صنعت مبدل شده است. از سوی دیگر، انجام داده‌کاوی نیازمند دسترسی به داده‌های گذشته است که با توجه به وجود ۱۷۶ پروژه مشارکت عمومی- خصوصی در این صنعت، می‌توان با سهولت به بررسی روند گذشته پرداخت.

درک داده‌ها: پس از جمع‌آوری داده‌ها از منابع مختلف، فرمت داده، تعداد متغیرها، نوع متغیرهای پیوسته و گسسته و مفهوم متغیرها، بررسی شد. همچنین با استفاده از روش‌های آماری مشخصات اصلی داده‌ها استخراج شد. خروجی این مرحله معمولاً به‌صورت گزارشی است که در آن اشتباهات داده‌ها، مقدار ازدست‌رفته و وجود رکوردهای تکراری، مشخص می‌شود. منظور از اشتباهات داده‌ها، مقدار بی‌معنی و داده دورافتاده و پرت است که در نتایج تأثیر می‌گذارد.

آماده‌سازی داده‌ها: با استفاده از گزارش خروجی بررسی کیفیت داده‌ها در مرحله قبل، در صورت وجود داده‌های ازدست‌رفته، داده‌های نویز و پرت، سطرهای تکراری باید حذف و داده پاک‌سازی شده برای مدل‌سازی آماده شود.

مدل‌سازی: در این مرحله ابتدا چندین تکنیک دسته‌بندی یا خوشه‌بندی بر اساس نوع داده‌ها انتخاب شد. با توجه به مرحله پیش پردازش، ممکن است در این مرحله بر اساس پارامترهای ارزیابی، نیاز به بازگشت به عقب داشته باشیم و تکنیک‌های مختلف دسته‌بندی و پارامترها را بررسی کنیم.

ارزیابی: در این مرحله می‌توان با توجه به نوع داده‌ها، شاخص‌های ارزیابی مدل را انتخاب کرد و بهترین روش و دسته‌بندی در هر مرحله بر اساس این ویژگی انتخاب می‌شود.

۳-۹- مدل پیشنهادی

با توجه به روش اجرای پژوهش، مدل پیشنهادی بر اساس تکنیک‌های داده‌کاوی، به شرح شکل ۱ است.

جمع‌آوری داده‌ها: به‌منظور به‌دست آوردن شاخص‌های کلیدی عملکرد مشارکت عمومی- خصوصی در صنعت آب و فاضلاب کشور ایران، در ابتدا اهداف کلیدی مشارکت بخش خصوصی در صنعت آب و فاضلاب مطابق با پژوهش‌های بوهل به‌صورت دانش و تجربیات فنی، اثربخشی عملیاتی، پاسخگویی به مشتریان، کارایی اقتصادی، تأمین مالی و جذب سرمایه‌های بخش خصوصی، دانش و

[1] http://invest.nww.ir

www.SID.ir

Archive of SID
dx.doi.org/ 10.22093/wwj.2020.211331.2970

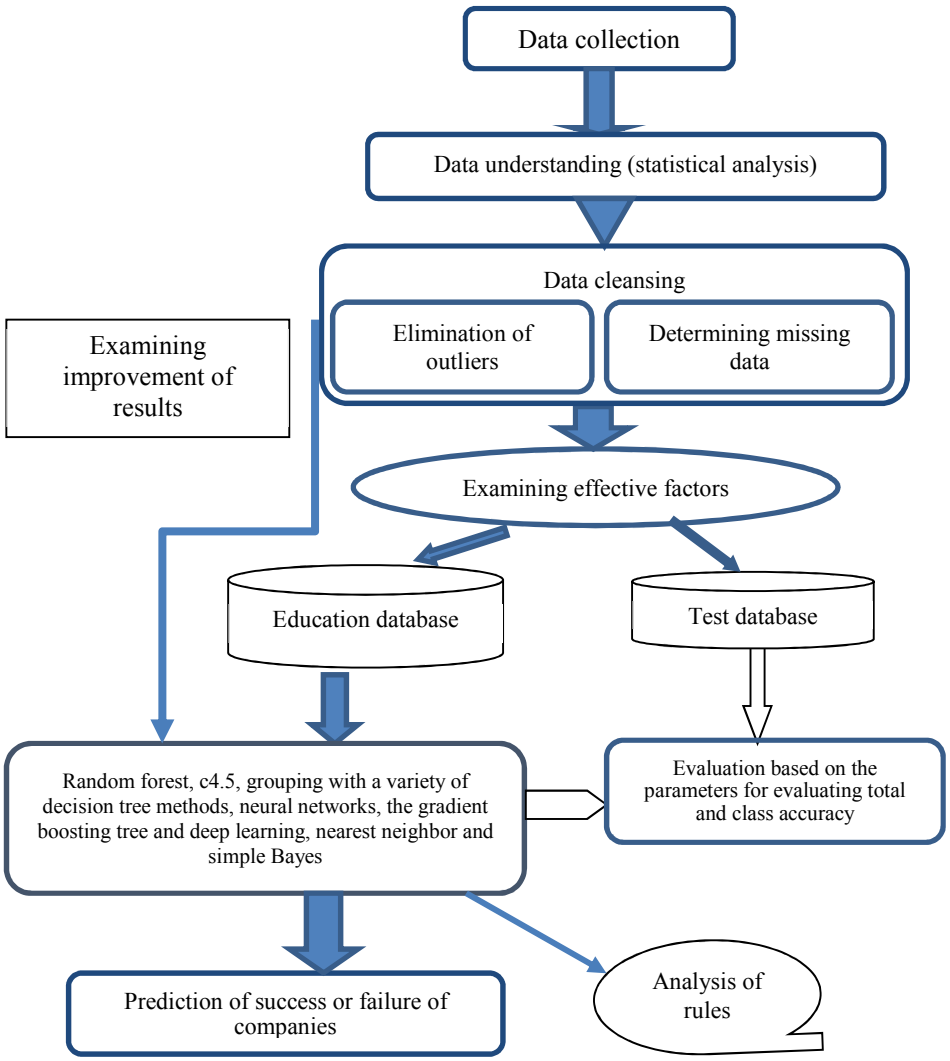

Fig. 1. The general model of the research

شکل ۱- مدل کلی پژوهش

تجربیات مدیریتی و مصون ماندن از مداخلات سیاسی، در نظر گرفته شد (Böhl, 2007). در ادامه، بر مبنای شاخص‌های مشارکت عمومی- خصوصی استخراج شده در بخش مرور ادبیات، شاخص‌های لازم برای ارزیابی تحقق هر یک از اهداف، مورد سنجش قرار گرفت و پس از تجمیع آنها، مجدداً نظرسنجی از خبرگان انجام شده است تا بررسی شود که اولاً همه ابعاد و شاخص‌ها به‌طور کامل در نظر گرفته شده باشد و ثانیاً حتی‌الامکان از دو بار لحاظ کردن شاخص‌ها اجتناب شود و در نتیجه شاخص‌هایی که هم‌پوشانی داشته و یا مشابه هستند حذف شده و در نهایت شاخص‌ها نهایی شد.

خبرگان صنعت به‌گونه‌ای انتخاب شده‌اند که بتوانند تمامی حوزه‌های مرتبط با مشارکت عمومی- خصوصی را بررسی کنند. در این راستا از تجربیات و نظرات ۸ خبره با ترکیب ۲ نماینده از جامعه سرمایه‌گذاران در پروژه‌های آب و فاضلاب، ۲ نماینده از مدیران عامل شرکت‌های آب و فاضلاب شهری، ۲ نماینده از شرکت مهندسی آب و فاضلاب کشور، ۱ نماینده از جامعه مشاورین و ۱ نفر نماینده از سازمان برنامه و بودجه کل کشور، استفاده شد.

شاخص‌های استخراج شده شامل قالب قراردادی، سال اجرای طرح، قابلیت فروش محصول در استان، تصفیه فاضلاب (درصد)، آب به حساب نیامده (درصد)، نسبت حوادث شبکه به طول کل شبکه (درصد)، نسبت حوادث انشعابات به کل انشعابات (درصد)،

www.SID.ir

dx.doi.org/10.22092/wwj.2020.211331.2970
Archive of SID

شکایات مربوط به کیفیت خدمات‌رسانی، کیفیت آب مصرفی (درصد)، جمعیت تحت پوشش خدمات آب (درصد)، جمعیت تحت میزان پوشش خدمات فاضلاب (درصد)، قطعی آب نسبت به تعداد انشعابات (درصد)، کمبود فشار آب (درصد)، تعرفه‌های متعارف (فروش به قیمت تمام شده در نقطه سربه‌سر)، سود (زیان) عملیاتی (درصد)، دوره وصول مطالبات، نسبت کاری (نسبت پوشش هزینه‌ها)، نسبت سرمایه‌گذاری طی سال به فروش (درصد)، نسبت گردش دارایی‌های ثابت (درصد)، سطح بدهکاری (نسبت پوشش بدهی)، سطح اندازه‌گیری (درصد)، هزینه تمام شده آب در نقطه سربه‌سر (ریال)، هزینه تمام شده فاضلاب دفع شده در نقطه سربه‌سر (ریال)، بهره‌وری نیروی انسانی (ناخالص)، اثربخشی نیروی انسانی (ناخالص)، سرانه هزینه نیروی انسانی (میلیون ریال)، درجه حرفه‌ای شدن شرکت، آموزش و غیبت بودند.

همان گونه که مشاهده می‌شود، شاخص‌ها شامل طیف وسیعی از موارد است که برای تحقق اهداف مشارکت عمومی- خصوصی موردنیاز است و کلیه بخش‌های مشارکت شامل مقتضیات نوع پروژه، مکان اجرا، زمان آن و سایر موارد را شامل می‌شود.

در ادامه و به‌منظور استخراج لیست پروژه‌ها، تمامی استان‌های کل کشور بررسی شده و پروژه‌های مشارکتی که در این استان‌ها با روش‌های اشاره شده محقق شده است، استخراج شد. جداول ارسالی به شرکت‌های آب و فاضلاب، شامل لیست پروژه‌ها و شاخص‌هایی بود که باید برای آنها ورود اطلاعات انجام می‌شد. با توجه به اولویت این پروژه در بخش آب و فاضلاب، اطلاعات تکمیل شده در هر استان، پس از تأیید مدیرعامل محترم آن استان به شرکت مهندسی آب و فاضلاب کشور ارسال شد و به‌منظور بررسی و صحت‌سنجی اولیه اطلاعات تکمیل شده، مجدداً اطلاعات شاخص‌ها با بانک داده مجامع شرکت‌ها که در شرکت مهندسی آب و فاضلاب کشور قرار دارد، تطبیق داده شد.

در ادامه، درصد موفقیت پروژه‌ها به‌صورت دقیق مشخص شد. درخصوص اعلام میزان موفقیت هر یک از پروژه‌های موجود، رقم دقیق موفقیت در بازه بین صفر تا ۱۰۰ توسط خبرگان موجود در ذی‌نفعان اصلی هر پروژه اعم از سرمایه‌گذار، سرمایه‌پذیر، شرکت مهندسی آب و فاضلاب کشور و مشاور طرح ارائه شد و سپس میانگین نظرات آنان استخراج شده و برای استفاده در مدل داده‌کاوی در سه دسته موفق، نیمه‌موفق و ناموفق تقسیم‌بندی شد.

تصمیم‌گیری در خصوص این که اعداد در چه بازه‌ای در هر یک از گروه‌ها قرار بگیرند نیز توسط خبرگان انجام شد. با توجه به نظر خبرگان برای ورود به الگوریتم اعداد میزان موفقیت در ۳ گروه Success موفقیت بالای ۷۵ درصد، semi-success موفقیت بین ۵۰ تا ۷۵ درصد و Failure موفقیت زیر ۵۰ درصد طبقه‌بندی شدند.

شناخت داده‌ها: در این مجموعه داده‌ها، ۱۷۶ پروژه با ۲۹ شاخص که در ۴ قالب قراردادی BOT[1]، BOO[2]، بیع متقابل و فاینانس جاری اجرا می‌شوند، مشخص شده است. آمار پروژه‌ها در قالب‌های قراردادی استفاده شده این صنعت، به‌صورت جدول ۱ است.

درخصوص شاخص‌های به‌دست آمده نیز، به غیر از قالب قراردادی، بقیه متغیرها به‌صورت عددی هستند. نام پروژه به‌عنوان ID و میزان موفقیت پروژه نیز متغیر هدف تعریف شد.

تحلیل آماری: اطلاعات آماری مجموعه داده برای تعدادی از شاخص‌ها به‌عنوان نمونه در جدول ۲ آمده است. ستون اول نوع داده، ستون دوم تعداد داده‌های ازدست‌رفته هر متغیر، ستون سوم و چهارم ماکسیمم و مینیمم متغیرهای عددی و کمترین و بیشترین مقدار متغیرهای گسسته و ستون آخر مقدار میانگین و انحراف معیار متغیرهای عددی و مقدار مد متغیرهای گسسته را نشان می‌دهد. نتایج حاصله نشان داد که ۹۰ درصد شاخص‌ها از نوع عددی هستند و هیچ یک از سطرهای داده مقدار ازدست‌رفته ندارند. علاوه بر این، با توجه به این که در برخی از شاخص‌ها انحراف معیار زیادی، مشاهده شد، پیش‌بینی شد که داده‌های پرت وجود دارد.

مصورسازی داده‌ها: در مرحله مصورسازی، نمودارهای مختلفی برای تحلیل اولیه داده‌ها ترسیم و بررسی شد. بررسی اولیه بر روی اطلاعات آماری، احتمال وجود داده‌های پرت را نشان داد. برای اطمینان از این مطلب، برای هر یک از شاخص‌ها نمودار quartile رسم شد. در این نوع از نمودارها که برای تشخیص داده‌های دورافتاده استفاده می‌شود، حباب‌ها، نقاط پرت را مشخص می‌کنند. نمودار quartile مربوط به شاخص‌های متعدد نشان داد که در اکثر شاخص‌ها داده پرت وجود دارد.

[1] Build-Operate-Transfer(BOT)
[2] Build-Operate-Own (BOO)

www.SID.ir

جدول ۱- توزیع قراردادهای مشارکت بخش خصوصی بر اساس قالب قراردادی

Table 1. Distribution of projects in the water and wastewater sector

Wastewater contracts	Treatment plant & pumping station	BOT contract	18 Contracts
		Buyback contract	31 Contracts
	Network	Current finance	73 Contracts
Water contracts	Treatment plant & transmission line	BOT contract	6 Contracts
	Desalination projects	BOO contract	48 Contracts

جدول ۲- اطلاعات آماری مجموعه داده

Table 2. Statistical data sets

Indicator	Type	Missing value	Statistical information		
Project success rate	Polynomial	0	Values: Success & Semi success & Failure	Values: Success & Semi success & Failure	Values: Success & Semi success & Failure
Year of project implementation	Integer	0	Average 2012	Average 2012	Average 2012
Ability to sell the product in the province	Integer	0	Average 4.1	Average 4.1	Average 4.1
Wastewater treatment (%)	Integer	0	Average 33.48	Average 33.48	Average 33.48
Non-revenue water (%)	Integer	0	Average 24.6	Average 24.6	Average 24.6
The ratio of network accidents to the network length (%)	Integer	0	Average 2.89	Average 2.89	Average 2.89
The ratio of connections accidents to the total connections (%)	Integer	0	Average 12.37	Average 12.37	Average 12.37

پاک‌سازی داده‌ها (تشخیص داده پرت): با توجه به این که بر اساس واریانس‌ها و نمودارهای مشخص شده در مبحث داده‌های ورودی، برخی از داده‌ها پرت هستند، حذف این داده‌ها می‌تواند به بهبود مدل کمک کند. داده پرت مشاهده‌ای است که در فاصله دورتری از سایر داده‌ها قرار می‌گیرد و با مقدار مورد انتظاری که داریم متفاوت است.

بر روی مجموعه داده دو روش LOF[1] و COF[2] برای تشخیص نقاط پرت و سپس حذف آن اعمال شد. این روش‌ها چگالی هر نقطه را محاسبه کرده و رتبه داده پرت را گزارش می‌دهند.

الگوریتم فاکتور داده پرت محلی یا LOF اولین بار در پژوهشی در سال ۲۰۰۰ معرفی شد. در این روش پرکاربرد، در مرحله اول برای هر نقطه K تا نزدیک‌ترین همسایه آن محاسبه می‌شود. سپس برای محاسبه فاصله قابل‌دسترسی یک نقطه نسبت به سایر نقاط، بیشترین مقدار فاصله دو نقطه و K تا نزدیک‌ترین همسایه‌اش در نظر گرفته می‌شود. در مرحله بعد برای چگالی محلی هر نقطه، معکوس میانگین فاصله قابل‌دسترسی نسبت به سایر نقاط محاسبه می‌شود و در آخر، رتبه داده پرت میانگین نرخ چگالی محلی هر نمونه داده نسبت به همسایگانش معرفی می‌شود. این الگوریتم نرخ داده پرت را به‌طور تقریبی بیش از ۱ و ۱/۱ معرفی می‌کند، یعنی داده‌های نرمال را کمتر از ۱/۱ و ۱ تعریف می‌کنیم. ازاین‌رو در نرم‌افزار بین انتخاب LOF outlier score<= ۱/۱ و LOF outlier score<=۱/۱۵ نتایج شبیه‌سازی شد و بهترین دامنه مقدار، پس از چندین مرتبه شبیه‌سازی، برابر ۱/۱۵ در نظر گرفته شد و داده‌هایی با وزن کمتر از ۱/۱۵ به‌عنوان داده نرمال در نظر گرفته و مابقی حذف شدند. پس از اعمال این الگوریتم، از مجموع ۱۷۶ پروژه، ۱۱۹ پروژه باقی ماند.

[1] Local Outlier Factor (LOF)
[2] Connectivity- based Outlier Factor (COF)

www.SID.ir

dx.doi.org/10.22092/ajwwj.2020.211331.2970
Archive of SID

الگوریتم COF، برای بهبود روش LOF پیشنهاد شده است. در اجرای این الگوریتم، نرخ داده پرت را به‌طور تقریبی از مقدار ۱ تا ۱/۱۵ شبیه‌سازی کرده و بهترین حد آستانه در مدل، نقاط پرت ۱/۱۵ به‌دست آمد. پس از اعمال این الگوریتم، از مجموع ۱۷۶ پروژه، ۱۲۶ پروژه باقی ماند.

نتایج اجرای روش‌های مذکور با مقدار k=10، مقدار نزدیکترین همسایه در جدول ۳ گزارش شده است که دامنه رتبه داده پرت، وجود نمونه‌های دور افتاده که باید حذف شوند را اثبات می‌کند. روش COF تعداد کمتری از نمونه‌ها را پرت تشخیص داده و حذف می‌کند.

در مرحله دسته‌بندی تفاوت دقت هر دسته‌بند بدون حذف داده پرت و با حذف داده پرت مشخص خواهد شد و نتایج دو روش LOF و COF مقایسه شد.

جدول۳- دامنه نمونه‌های پرت در هر روش

Table 3. Outlier range rate in any method

Method	Setting	Output data variable: true numeric value
LOF	Small K: 10 Capital K: 20 Distance: Euclidean Type of measurement: Combined measurement	Output data range: 1.192-0.983 Average: 1.113 SD: 0056
COF	Small K: 10 Distance: Euclidean Type of measurement: Combined measurement	Output data range: 0.469-2.765 Average: 1.110 SD: 0.432

انتخاب ویژگی: در این بخش، روش انتخاب ویژگی شاخص آنتروپی (سود اطلاعاتی[1]) برای ارزیابی و انتخاب متغیرهای تأثیرگذار انتخاب شد. این روش وزن تأثیری ویژگی‌ها بر اساس سود اطلاعاتی هر شاخص را در طبقه‌بندی شکست و یا موفقیت پروژه، با مقداری بین صفر و یک در نرم‌افزار رپیدماینر مطابق با شکل ۲ گزارش داد که بر اساس آن نتایج زیر حاصل شد:

با توجه به وزن ویژگی‌ها، ۹۸ درصد ویژگی‌ها وزن زیادی داشتند و در طبقه‌بندی و پیش‌بینی شکست و یا موفقیت پروژه‌ها، تأثیر بیشتری داشتند.

سطح شکایات مربوط به کیفیت خدمات‌رسانی، به‌طور مستقیم با بیشترین وزن ۱ روی پیش‌بینی شکست تأثیر داشت.

نوع قالب قراردادی، نسبت پوشش هزینه‌ها و تعرفه‌های متفاوت تأثیرگذارترین متغیرها بود.

بحث تصفیه فاضلاب، درجه حرفه‌ای شدن و سطح بدهکاری با وزن پایین، تأثیر بسیار کمی در پیش‌بینی شکست و یا موفقیت پروژه‌ها دارد.

۳-۱۰- مدل‌سازی (طبقه‌بندی)

در این پژوهش و در گام دسته‌بندی، پس از انجام مراحل پیش‌پردازش داده‌ها، مدل‌سازی با الگوریتم‌های مختلف انجام شد. نتایج دسته‌بندها یک بار بدون انتخاب ویژگی و حذف داده پرت، یک بار با روش LOF و انتخاب ویژگی سود اطلاعاتی و یک بار با روش COF و انتخاب ویژگی پیاده‌سازی و نتایج مقایسه و تحلیل شد.

برای اعتبارسنجی و تقسیم داده‌ها به دو قسمت آموزشی و آزمایشی و به‌منظور آموزش بهتر داده‌ها، از روش اعتبارسنجی ضربدری به‌جای اعتبارسنجی ساده استفاده شد. در این نوع اعتبارسنجی، داده‌ها به K زیرمجموعه افراز می‌شوند. از این K زیرمجموعه، هر بار یکی برای اعتبارسنجی و K-1 تای دیگر برای آموزش به‌کار می‌روند. این روال K بار تکرار می‌شود و همه داده‌ها دقیقاً یک بار برای آموزش و یک بار برای اعتبارسنجی به‌کار می‌روند. در نهایت میانگین نتیجه این K بار اعتبارسنجی، به‌عنوان یک تخمین نهایی برگزیده می‌شود.

پس از مراحل پیش‌پردازش مجموعه داده‌ها، از رده‌بندهای پایه مانند درخت تصمیم، بیزساده، شبکه عصبی و در ادامه از روش‌های ترکیبی مانند یادگیری عمیق، آدابوست درخت تصمیم، جنگل تصادفی، گرادیان بوستینگ و یک روش ترکیبی شبکه عصبی و درخت تصمیم با نام استکینگ، در نرم‌افزار استفاده شد. پارامترهای در نظر گرفته شده بر اساس چندین مرحله پیاده‌سازی و بررسی نتایج، برای درخت تصمیم شاخص سود اطلاعاتی، حداکثر عمق ۱۵، حداقل سایز برگ ۳ و حداقل تعداد رکورد جداسازی ۴ بود. شبکه عصبی پرسپترون با ضریب مومنتوم ۰/۸ و نرخ یادگیری ۰/۳ تعیین شد. گرادیان بوستینگ و جنگل تصادفی با ۲۰ درخت در نظر گرفته شد.

[1] Information Gain

www.SID.ir

Archive of SID
dx.doi.org/ 10.22093/wwj.2020.211331.2970

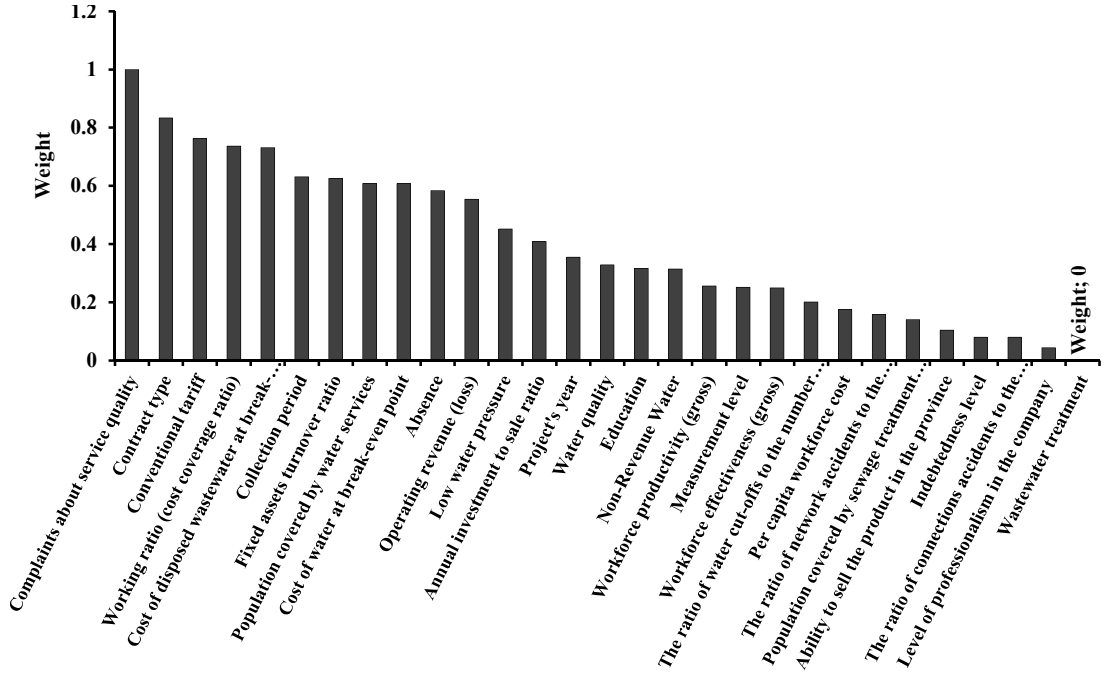

Fig. 2. Weight of Indicator

شکل ۲- وزن‌دهی شاخص‌ها

از مجموع ۲۹ شاخص، پس از شبیه‌سازی ۲۷ شاخص انتخاب و شاخص‌های تصفیه فاضلاب و درجه حرفه‌ای شدن شرکت که وزن بسیار کمی داشتند حذف شدند، زیرا حذف این دو شاخص دقت دسته‌بندها را افزایش داد. پس از شبیه‌سازی، نتایج بر اساس پارامترهای ارزیابی مقایسه شد.

۳-۱۱- پارامترهای ارزیابی

برای ارزیابی الگوریتم‌ها از معیارهایی از جمله دقت، استفاده شده است. قبل از آن که به بررسی انواع مهم معیارهای دسته‌بندی بپردازیم، لازم است که مفهوم ماتریس درهم‌ریختگی که در این بخش استفاده شده، تشریح شود. این ماتریس چگونگی عملکرد الگوریتم دسته‌بندی را با توجه به مجموعه داده ورودی به تفکیک انواع رده‌های مسئله دسته‌بندی نشان می‌دهد. با توجه به جدول ۴، مفاهیم مثبت درست[1]، مثبت غلط[2]، منفی درست[3] و منفی غلط[4] به شرح زیر است:

- مثبت درست: این مقدار بیانگر تعداد رکوردهایی است که دسته واقعی آنها مثبت بوده و دسته‌بند نیز رده آنها را به‌درستی مثبت تشخیص داده است.
- مثبت غلط: این مقدار بیانگر تعداد رکوردهایی است که دسته واقعی آنها منفی بوده و دسته‌بند رده آنها را به نادرستی مثبت تشخیص داده است.
- منفی درست: این مقدار بیانگر تعداد رکوردهایی است که دسته واقعی آنها منفی بوده و دسته‌بند نیز دسته آنها را به‌درستی منفی تشخیص داده است.

جدول ۴- ماتریس درهم‌ریختگی

Table 4. Clutter matrix

		Estimated samples	
		Positive batch (+)	Negative batch (-)
Real samples	Negative Batch (-)	FP	TN
	Positive Batch (+)	TP	FN

[1] True Positive-TP
[2] False Positive_FP
[3] True Negative-TN
[4] False Negative_FN

www.SID.ir

Archive of SID
dx.doi.org/10.22093/wwj.2020.211331.2970

- منفی غلط: این مقدار بیانگر تعداد رکوردهایی است که دسته واقعی آنها مثبت بوده و دسته‌بند رده آنها را به نادرستی منفی تشخیص داده است.

اولین معیار، دقت یا نرخ دسته‌بندی است و به این مفهوم است که دسته‌بندی، چند درصد رکوردهای آزمایشی را به‌درستی تشخیص داده است. این دقت بر اساس مفاهیم ماتریس با توجه به معادله ۲ قابل‌محاسبه است

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (۲)$$

معیاردیگر، Recall یا دقت کلاس‌های مثبت و یا منفی است. در این معیار صورت کسر تعداد تشخیص‌های درست کلاس مثبت (منفی) نسبت به کل نمونه‌هایی که واقعاً مثبت (منفی) بوده‌اند

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

$$\text{Recall}- = \frac{TP}{FN + TP} \quad (۴)$$

۴- نتایج و بحث

بر مبنای اصول گفته شده در بخش‌های قبل، مدل‌سازی انجام شده و نتایج پیاده‌سازی تکنیک‌های دسته‌بندی در سه مرحله برای هر دسته‌بند اجرا شد که نتایج دقت کل و دقت هر کلاس نوع پروژه، به شرح شکل‌های ۳، ۴ و ۵ است.

با توجه به شکل ۳، دقت کلاس شکست در تمام دسته‌بندها به‌دلیل نمونه پروژه‌های کم شکست (عدم‌توازن در کلاس) بسیار کمتر از کلاس‌های دیگر است، اما به‌طور کل در هر سه دقت در همه دسته‌بندها نتایج نسبتاً متوسطی داریم. روش‌های جنگل تصادفی و نایوبیز دقت بیشتری نسبت به سایر دسته‌بندها دارند.

با توجه به شکل ۴، دقت کل مدل‌ها با حذف نمونه‌های پرت، افزایش داشته است. در روش‌های درختی تا میانگین ۷ درصد با حذف نمونه‌های دورافتاده افزایش دقت داشتیم.

الگوریتم‌های پایه مانند شبکه عصبی و درخت تصمیم در کلاس failure نتایج ضعیفی دارند، اما روش‌های ترکیبی آدابوست، جنگل تصادفی، گرادیان و استکینگ به‌دلیل تقسیم داده و متوازن شدن کلاس‌ها دقت بیشتری گزارش دادند. در مجموع بیشترین دقت با جنگل تصادفی در هر سه کلاس دقت، گزارش شد.

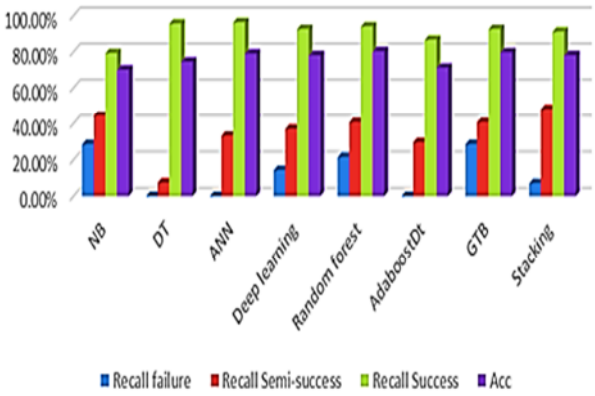

Fig. 3. Results of models accuracy without feature selection and deletion of outliers

شکل ۳- نتایج دقت مدل‌ها بدون انتخاب ویژگی و حذف داده پرت

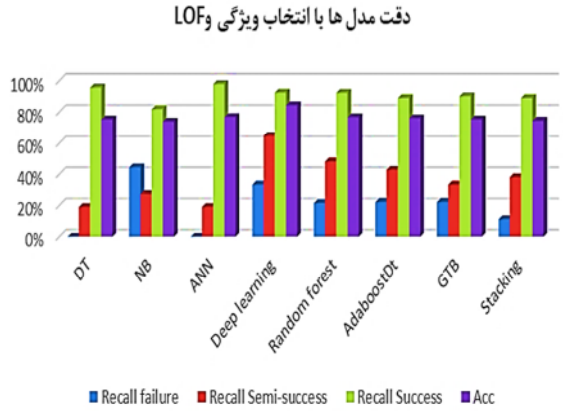

Fig. 4. Results of models accuracy with feature selection and LOF outlier data deletion

شکل ۴- نتایج دقت مدل‌ها با انتخاب ویژگی و حذف داده پرت LOF

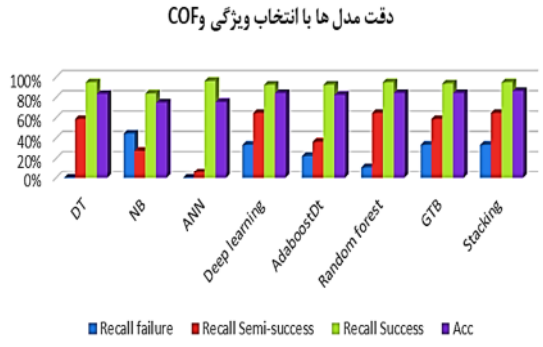

Fig. 5. Results of models accuracy with feature selection and COF outlier data deletion

شکل ۵- نتایج دقت مدل‌ها با انتخاب ویژگی و حذف داده پرت COF

www.SID.ir

Archive of SID
dx.doi.org/ 10.22093/wwj.2020.211331.2970

جدول ۵- نتایج الگوریتم منتخب استکینگ در کلاس‌های مختلف
Table 5. The results of Stacking algorithm

	True success	True semi-success	True failure	Class precision
Pred. success	75	5	1	92.59%
Pred. semi-success	4	11	3	61.11%
Pred. failure	0	1	2	66.67%
Class recall	94.94%	64.71%	33.33%	

accuracy: 86.27% +/-10.14% (micro average: 86.27%)

همان گونه که در جدول ۵ مشاهده می‌شود، بیشترین دقت کل برای الگوریتم استکینگ با روش حذف داده پرت COF و به مقدار ۸۶/۲۷ درصد گزارش شد.

این مدل همچنین در مقایسه با سایر مدل‌ها در کلاس‌های موفقیت، شکست و نیمه‌موفق نیز دقت زیادی گزارش داد در صورتی که مدل‌های دیگر مانند شبکه عصبی و درخت تصمیم در کلاس‌های نیمه‌موفق و شکست به‌دلیل عدم توازن نمونه‌ها، دقت بسیار پایین و حتی در برخی موارد صفر درصد را گزارش دادند.روش حذف داده پرت COF، نتایج بسیار بهتری نسبت به LOF گزارش داد و در اکثر مدل‌ها دقت بیشتری داشت که نشان می‌دهد این روش داده‌های پرت را بهتر حذف کرده است.

مدلی موفق است که در همه کلاس‌های موفق، نیمه‌موفق و ناموفق، دقت نتایج خوبی داشته باشد، به دلیل آن که داده‌های ورودی در این پژوهش نامتوازن است (۹۰ درصد پروژه‌ها موفق هستند) مدل بر روی نمونه‌های کلاس اکثریت، دقت بیشتری دارند و دقت در کلاس شکست پایین است، اما در مقایسه مدل‌ها و با توجه به نمودارها، مدل‌های استکینگ، یادگیری عمیق و گرادیان بوستینگ در مقایسه با سایر مدل‌ها جزو روش‌های برتر این پژوهش بوده و دقت آنها در کلاس شکست بسیار بهتر از سایر روش‌ها است.دقت درخت تصمیم پایه با روش‌های ترکیبی آدابوست، جنگل تصادفی و استکینگ افزایش زیادی داشت. این روش‌ها که از تعدادی درخت استفاده می‌کنند، بر اساس رأی اکثریت مدل‌ها کلاس را پیش‌بینی کرده و به‌طورکلی دقت درخت را به‌خصوص روی داده‌های نامتوازن افزایش می‌دهند، زیرا بخشی از مجموعه داده را به هر درخت اعمال شده و نسبت کلاس‌ها در حالت ترکیبی متوازن‌تر از حالت یک روش درختی است. همچنین در روش‌های درختی مانند استکینگ، دقت در فاز دوم پیش‌بینی تقویت شده و افزایش می‌یابد.

۴-۱- تحلیل شاخص‌های تأثیرگذار

یکی از مواردی که از ابتدای پژوهش همواره مورد تأکید و درخواست صنعت آب و فاضلاب بوده است، رتبه‌بندی شاخص‌ها بوده تا از طریق بهبود شاخص‌های با اولویت بالاتر بتوانند میزان موفقیت پروژه‌های مشارکتی را به‌طور مؤثری بهبود بخشند. رتبه‌بندی شاخص‌ها از نظر میزان تأثیرگذاری بر موفقیت پروژه‌های مشارکت عمومی- خصوصی، بر مبنای روش شاخص سود اطلاعاتی(آنتروپی) انجام شده که ۱۵ شاخص اول مؤثر در مشارکت عمومی- خصوصی به ترتیب اولویت درجدول ۶ نشان داده شده است.

همان گونه که مشاهده می‌شود، در صنعت آب و فاضلاب ایران، موفقیت پروژه‌های مشارکت عمومی- خصوصی بیشترین تأثیر را از شاخص "شکایات مربوط به کیفیت خدمات‌رسانی" می‌پذیرد. با توجه به این که بیش از ۴۰ درصد از پروژه‌های موجود به‌صورت فاینانس جاری و با مشارکت مردم در تأمین مالی طرح انجام شده است، این مطلب گویای تأثیر رضایت مردم از کیفیت خدمات بر روی موفقیت و استمرار پروژه است. بر مبنای اطلاعات حاصل از جدول ۶، در بین ۱۵ شاخص کلیدی موفقیت پروژه‌های مشارکت عمومی- خصوصی در صنعت آب و فاضلاب، ۸ شاخص آن به‌طور مستقیم با مسائل مربوط به تأمین مالی طرح‌ها در ارتباط هستند. این امر بیانگر این است که در صنعت آب و فاضلاب ایران، برخلاف سایر کشورهای توسعه یافته که حرکت به سمت اجرای مشارکتی طرح‌ها به‌دلیل افزایش بهره‌وری، کیفی‌سازی طرح‌ها و استفاده از دانش و تجربیات بخش خصوصی است، همچنان دغدغه اصلی تامین مالی طرح‌ها است.

www.SID.ir

DOI: 10.22093/wwj.2020.211331.2970
Archive of SID

جدول ۶- ۱۵ شاخص کلیدی موفقیت پروژه‌های مشارکت عمومی- خصوصی در صنعت آب و فاضلاب

Table 6. 15 Key performance indicators of success in public-private partnership projects in the water and wastewater industry

Priority	Name of indicator
1	Complaints about service quality
2	Contract format
3	Different tariffs (sale price)
4	Working ratio (cost coverage ratio)
5	Cost of sewage
6	Collection period
7	Fixed assets turnover ratio
8	Cost of water at break-even point
9	Population covered by water services
10	Absence
11	Operating profit and loss (in percentage)
12	Low water pressure (in percentage)
13	Annual investment to sale ratio
14	Year of project implementation
15	Water quality (in percentage)

بر مبنای نتایج حاصل شده، قالب قراردادی به‌عنوان دومین شاخص مهم و تأثیرگذار در موفقیت پروژه مشارکتی مطرح است و در نتیجه، شرکت‌های آب و فاضلاب به راحتی می‌توانند قبل از شروع به عملیات اجرایی، نتایج حاصل از اجرای پروژه در هر یک از قالب‌ها را ارزیابی و بررسی کرده و مناسب‌ترین قالب قراردادی برای اجرای پروژه را انتخاب کنند. تعرفه‌های متعارف سومین شاخص تأثیرگذار در موفقیت یک پروژه مشارکتی است. این امر نشان می‌دهد که در صورت حذف تعرفه‌های تکلیفی و از طریق ایجاد فضای رقابتی بر مبنای هزینه‌های واقعی تولید آب و پساب، می‌توان بسیاری از پروژه‌های بخش آب و فاضلاب را به‌صورت خودگردان و با مشارکت بخش خصوصی تکمیل کرده و این صنعت تا حد زیادی از وابستگی به منابع دولتی دور شود.

۴-۲- تحلیل قوانین

با توجه به مدل‌های درختی گرادیان بوستینگ، جنگل تصادفی و استکینگ که در این پژوهش بر روی داده‌ها دقت زیادی دارند (نمونه‌ای از خروجی در شکل ۶ آمده است)، در این بخش به بررسی مهم‌ترین قوانین قابل‌تولید و الگوهای پنهان موجود در داده‌ها پرداخته شده است. البته بیشتر قوانین گزارش شده، آن دسته از قوانینی هستند که احتمال وقوع بیشتر و خطای کمتری داشته و در نتیجه مفیدتر هستند.

قوانین استخراج شده به صورت زیر هستند:

۱- "اگر تعرفه‌های متعارف بیشتر از ۱۹/۳ درصد باشد و شکایت کمتر از ۳۸/۵ باشد، پروژه با احتمال زیاد موفق می‌شود". این امر نشان‌دهنده اهمیت زیاد این دو شاخص بوده که در صورتی که شرکت‌های آب و فاضلاب بتوانند در بهبود آنها قدم بردارند، در کوتاه‌ترین زمان خواهند توانست به موفقیت پروژه‌ها نزدیک شوند. این قانون شامل دو فاکتور اساسی است که یکی از آنها تعرفه‌های متعارف برای خدمات صنعت است و با توجه به تکلیفی بودن این تعرفه‌ها، نیازمند سیاست‌گذاری در خصوص بهبود آنها هستیم.

بخش دیگر افزایش رضایت مشترکین است. در واقع در صورتی که بخش آب و فاضلاب بتواند از طریق بررسی دقیق سرمایه‌گذاران در زمان مناقصات و مزایدات، رضایت مشترکان را جلب کند، احتمال موفقیت پروژه مشارکت عمومی- خصوصی تا حد زیادی افزایش می‌یابد.

۲- "اگر سود و زیان عملیاتی بیشتر از ۱۴۵/۹ و قالب قراردادی بیع‌متقابل باشد، احتمالاً پروژه موفقی داریم". استخراج این قانون بیانگر نقش شاخص سود و زیان عملیاتی در موفقیت یا عدم

www.SID.ir

Archive of SID
dx.doi.org/ 10.22093/wwj.2020.211331.2970

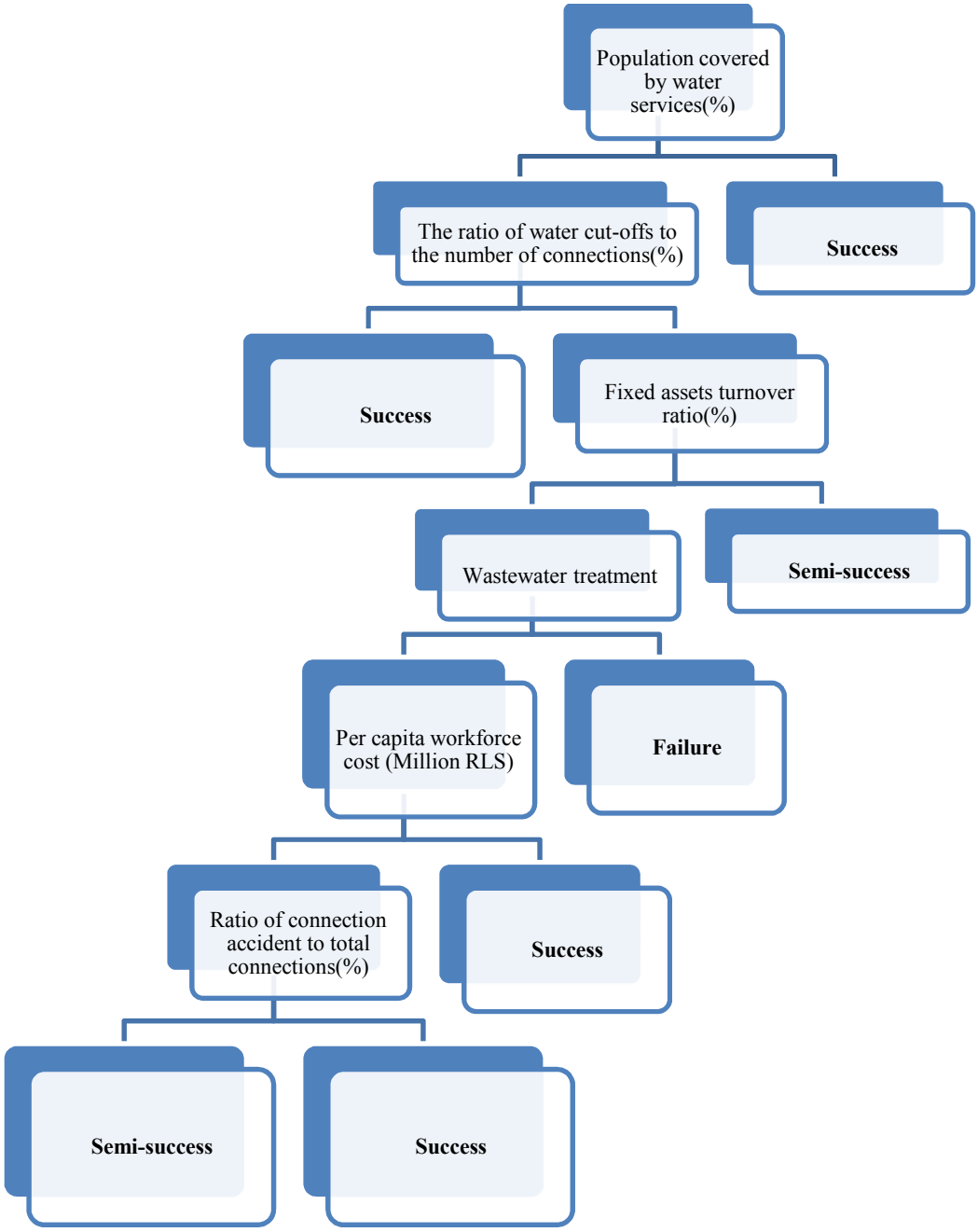

Fig. 6. Sample tree

شکل ۶- مدل درخت روش استکینگ

موفقیت پروژه‌های بیع‌متقابل است. با توجه به استقرار پروژه‌های بیع‌متقابل در فلات مرکزی ایران، شرکت‌های آب و فاضلاب مستقر در این استان‌ها باید تلاش بیشتری در زمینه بهبود این شاخص داشته باشند.

۳- "اگر غیبت کمتر از ۷۹/۵ روز باشد، احتمالاً در پروژه شکست داریم". شاخص غیبت در واقع شاخص شماره ۱۰ در رتبه‌بندی به دست آمده بوده و تنها شاخصی است که با عوامل انسانی موجود در بخش آب و فاضلاب در ارتباط است. بر مبنای بررسی‌های به عمل آمده، یکی از دلایل استقبال بخش خصوصی از مشارکت در طرح‌های صنعت آب و فاضلاب، حضور و پاسخ‌گویی نیروهای سازمانی این صنعت است.

با توجه به ساختار سازمانی ایجاد شده در بخش آب و فاضلاب

www.SID.ir

DOI: 10.52547/wwj.2020.211331.2970
Archive of SID

و اختصاص دفتر جداگانه در زیر مجموعه مدیرعامل برای تجهیز منابع مالی و توسعه مشارکت بخش غیردولتی، درخواست‌های سرمایه‌گذاران به خوبی در این صنعت مورد توجه قرار گرفته و در نتیجه می‌تواند پروژه‌های با موفقیت بیشتری را به همراه داشته باشد.

۴- "اگر دوره مطالبات کمتر از ۹۰/۵ روز باشد، پروژه احتمالاً نیمه‌موفق است." از جمله عوامل بسیار مؤثر در بقا و تداوم یک پروژه مشارکت عمومی- خصوصی، تأمین جریان نقدینگی سرمایه‌گذار است. با توجه به وجود سرمایه در گردش و توان مالی سرمایه‌گذاران موجود در صنعت آب و فاضلاب، ضروری است این بخش برنامه‌ریزی‌های خود را به‌گونه‌ای انجام دهد که مطالبات معوق بخش خصوصی، بیشتر از ۳ ماه نباشد.

۵- "اگر دوره وصول مطالبات کمتر از ۵۳/۵ روز باشد، آنگاه با احتمال زیاد پروژه موفق داریم". این قانون در کنار قانون شماره ۹، ضرورت پرداخت معوقات بخش خصوصی در بازه ۲ تا ۳ ماه را نشان می‌دهد.

۴-۳- تحلیل قالب مناسب قرارداد در استان‌های مختلف

در این مرحله با استفاده از نمودار Bar Stacked در نرم‌افزار رپیدماینر، ارتباط بین درصد موفقیت و یا شکست پروژه‌ها در استان‌های مختلف بر اساس نوع قالب قراردادی بررسی شد و نتایج زیر به‌دست آمد:

۱- استان‌هایی که در مرکز ایران واقع شده‌اند، در اجرای پروژه‌ها به روش بیع‌متقابل توفیق بیشتری داشته‌اند که این امر به‌دلیل استقرار صنعت در این مناطق و محدودیت منابع آبی در این شهرها است.

۲- اجرای پروژه‌ها به روش فاینانس جاری در برخی از استان‌ها از جمله تهران، خراسان رضوی، اصفهان و شیراز، با موفقیت بیشتری همراه بوده است. در واقع با توجه به این که این قالب قراردادی با مشارکت مردم انجام می‌شد، در مناطقی که تراکم جمعیتی بیشتر است، حجم سرمایه‌گذاری در بخش اجرای شبکه کاهش می‌یابد و سطح درآمد مردم به نسبت بیشتر است، تأمین مالی مناسبی در این پروژه‌ها انجام شده و در نتیجه موفقیت این پروژه‌ها با احتمال بیشتری رخ می‌دهد.

۳- اجرای پروژه‌ها به روش BOO در استان‌های نوار ساحلی کشور بسیار موفق‌تر از سایر استان‌ها است. در این استان‌ها قالب پروژه‌ها در حوزه آب‌شیرین‌کن و به روش BOO اجرا می‌شوند.

۵- نتیجه‌گیری

در این پژوهش، برای دسته‌بندی پروژه‌های شکست‌خورده و موفق و سپس ساختن مدل پیش‌بینی، ۱۷۶ پروژه شامل کلیه طرح‌های مشارکتی شرکت‌های آب و فاضلاب سراسر کشور جمع‌آوری شد و اطلاعات آنها با ۲۷ شاخص در ۴ قالب قراردادی BOT، BOO، بیع‌متقابل و فاینانس جاری از کل کشور تکمیل شد. برای درک داده‌ها، خصوصیات آماری، نمودار هیستوگرام و پراکندگی برخی متغیرها رسم و تحلیل شد. به‌دلیل انحراف معیار زیاد برخی شاخص‌ها، در بخش پیش‌پردازش از تکنیک‌های تشخیص داده پرت LOF و COF استفاده شد. در بخش دسته‌بندی الگوریتم‌های درخت تصمیم، یادگیری عمیق، شبکه عصبی چندلایه، بیزساده، آدابوست درخت تصمیم، جنگل تصادفی و گرادیان بوستینگ، ابتدا بدون انتخاب ویژگی و حذف داده پرت، سپس با حذف داده پرت و انتخاب متغیرهای تأثیرگذار، مدل‌سازی شده و در نرم‌افزار رپیدماینر پیاده‌سازی شد.

با توجه به نتایج به‌دست آمده، شاخص‌های مشارکت عمومی- خصوصی به‌ترتیب اولویت استخراج و ۱۵ شاخص برتر معرفی شد. بر مبنای یافته‌های مدل‌سازی، الگوریتم‌های آدابوست، استکینگ و جنگل تصادفی با دقت بیش از ۸۶ درصد، روش‌های برتر در پیاده‌سازی مدل بودند. روش حذف داده پرت COF به‌دلیل پاک‌سازی و تشخیص نمونه‌های پرت بیشتر، بهبود خوبی در دقت مدل‌های دسته‌بندی داشت و لزوم اجرای مرحله پاک‌سازی و پیش‌پردازش را در داده‌های سازمانی در پروژه‌های داده‌کاوی اثبات کرد. با توجه به این امر، تحلیل قوانین بر مبنای خروجی‌های درخت تصمیم و تحلیل نتایج بر اساس خروجی الگوریتم استکینگ انجام شد. بر مبنای نتایج به‌دست آمده از مدل، می‌توان توصیه‌های زیر را برای بهبود عملکرد پروژه‌های مشارکت عمومی- خصوصی در صنعت آب و فاضلاب مطرح کرد:

۱- با توجه به محدودیت‌های مالی و انسانی، پیشنهاد می‌شود تأکید اصلی بر بهبود شاخص‌هایی از جمله "شکایات مربوط به کیفیت خدمات‌رسانی" باشد که بتواند بیشترین تأثیر را در موفقیت پروژه‌های مشارکتی این صنعت داشته باشد.

www.SID.ir

Archive of SID
dx.doi.org/ 10.22093/wwj.2020.211331.2970

۲- بر مبنای نتایج حاصل از پژوهش، بهترین روش اجرای مشارکتی بر اساس نوع پروژه‌های موجود در صنعت آب و فاضلاب به‌صورت جدول ۷ است.

جدول ۷ - روش بهینه اجرای مشارکتی بر اساس نوع پروژه‌ها

Table 7. Optimal method of collaborative implementation by type of projects

Sewage contracts	**Construction of refineries and pumping stations**	**BOT/ Buyback**
	Construction of sewer networks	Buyback
Water sector contracts	Construction of treatment plants and transmission lines	BOT
	Construction of desalination plant and quality restoration	BOO

همان گونه که مشاهده می‌شود، قالب قراردادی بیع‌متقابل در پروژه‌های فاضلاب می‌تواند بسیار موفق باشد، ولی با توجه به تحلیل انجام شده درخصوص استان‌های مختلف و با توجه به این که کالای پساب هنوز ارزش واقعی خود را در تمامی نقاط کشور به‌دست نیاورده است، در حال حاضر بخش آب و فاضلاب این قالب قراردادی را در مناطق مرکزی که صنعت و مشتریان بالقوه وجود داشته و مشکلات کمبود آب نیز بیش از سایر مناطق است، استفاده می‌شود. در سایر نقاط، روش بهینه مشارکت برای اجرای طرح‌های فاضلاب، روش BOT است که به‌دلیل محدودیت منابع مالی دولت چندان مورد استقبال صنعت آب و فاضلاب نبوده و تمایل این بخش اجرای پروژه‌های تصفیه‌خانه به‌صورت BOT خودگردان است. در بخش احداث شبکه‌های فاضلاب، با توجه به مستندات قانونی جدید و امکان افزایش میزان حق انشعاب دریافتی از مردم با موافقت شورای شهر، می‌توان این پروژه‌ها را از طریق درآمدهای حاصل از این حوزه و در قالب قراردادی فاینانس جاری انجام داد.

۱- با توجه به اهمیت شاخص انتخاب قالب قراردادی مناسب، توصیه می‌شود که شرکت‌های آب و فاضلاب قبل از شروع اجرای پروژه و بر مبنای مدل ارائه شده، موفقیت حاصل از اجرای طرح در هر یک از قالب‌های قراردادی بررسی شده و بر مبنای نتایج آن، قالب قراردادی مناسب هر پروژه را انتخاب کنند.

۲- شاخص " تعرفه‌های متعارف" از جمله شاخص‌های مهم در موفقیت پروژه‌های مشارکتی است که تحقق آن باید مدنظر سیاست‌گذاران بخش آب و فاضلاب باشد.

۳- با توجه به اهمیت شاخص‌های مرتبط با نیروی انسانی در موفقیت پروژه‌های مشارکتی، حضور فعال و دائمی نیروی انسانی متخصص برای جلب اعتماد بخش خصوصی و موفقیت پروژه‌های مشارکتی، بسیار ضروری است.

۴- بر مبنای مدل‌سازی انجام شده و قوانین استخراج شده از آن، آستانه تحمل سرمایه‌گذاران در حوزه تأمین نقدینگی ۹۰ روز است و برنامه‌ریزی بخش آب و فاضلاب باید به‌گونه‌ای باشد که مطالبات سرمایه‌گذاران، قبل از این موعد پرداخت شود.

۶- قدردانی

این پژوهش برگرفته از رساله دکترا دانشکده مدیریت و حسابداری دانشگاه علامه طباطبایی است و بر خود لازم می‌دانیم از حمایت‌های مجموعه آب و فاضلاب در شرکت مهندسی آب و فاضلاب کشور و شرکت‌های استانی که در گردآوری اطلاعات ما را یاری رساندند، نهایت قدردانی را به عمل آوریم.

References

Akintoye, A., Hardcastle, C., Beck, M., Chinyio, E. & Asenova, D. 2003. Achieving best value in private finance initiative project procurement. *Construction Management and Economics*, 21, 461-470.

Al-Radaideh, Q. A. & Al Nagi, E. 2012. Using data mining techniques to build a classification model for predicting employees performance. *International Journal of Advanced Computer Science and Applications*, 3(2), 144-151.

Alegre, H. & Association, I. W. 2006. *Performance indicators for water supply services*, IWA publishing. London, UK.

www.SID.ir

Ali, M. & Qamar, A. M. 2013. Data analysis, quality indexing and prediction of water quality for the management of rawal watershed in Pakistan. *8th International Conference on Digital Information Management (ICDIM 2013)*, Islamabad, Pakistan, IEEE, 108-113.

Azhar, S. A. S., Johar, H., Baki, S. R. M. S. & Tahir, N. M. 2013. Optimization of water quality monitoring based on fuzzy algorithms. *IEEE Conference on Systems, Process and Control (ICSPC, 2013)*, IEEE, 283-288.

Beikzadeh, M. R., Phon-Amnuaisuk, S. & Delavari, N. 2008. Data mining application in higher learning institutions. *Informatics in Education-An International Journal*, 7(1), 31-54.

Böhl, C. G. P. 2007. *Development of a knowledge based decision support system for private sector participation in water and sanitation utilities*, Oldenbourg Industrieverlag, Munichen, Germany.

Chen, W. S. & Du, Y. K. 2009. Using neural networks and data mining techniques for the financial distress prediction model. *Expert Systems with Applications*, 36, 4075-4086.

Cheng, H., Lu, Y. C. & Sheu, C. 2009. An ontology-based business intelligence application in a financial knowledge management system. *Expert Systems with Applications*, 36, 3614-3622.

Company, N. W. A. W. 2009. *National wastewater financing and investment strategy*. Ministry of Energy, Tehran, Iran. (In Persian)

Cui, C., Liu, Y., Hope, A. & Wang, J. 2018. Review of studies on the public–private partnerships (PPP) for infrastructure projects. *International Journal of Project Management*, 36, 773-794.

Delmon, J. 2015. *Creating a framework for public-private partnership programs: a practical guide for decision-makers*, World Bank, Washington, DC, USA.

Freund, Y. & Schapire, R. E. 1995. A desicion-theoretic generalization of on-line learning and an application to boosting. *European Conference on Computational Learning Theory*, Bertinoro, Italy, Springer, 23-37.

Golabchi, M. & Nourzaei, E. A. 2015. Selecting the best PPP method in rail projects by using AHP methods. *Journal of Transportation Engineering*, 6(3), 523-537. (In Persian)

Hall, D., Lobina, E. & Motte, R. D. L. 2005. Public resistance to privatisation in water and energy. *Development in Practice*, 15, 286-301.

Hodge, G. A., Greve, C. & Boardman, A. 2010. *International handbook on public–private partnerships*, Edward Elgar Publishing, UK.

Jenabi, G. & Mirroshandel, S. A. 2014. Using data mining techniques for improving customer relationship management. *European Online Journal of Natural and Social Sciences: Proceedings*, 2, 3143-3149.

Kirimi, J. M. & Moturi, C. A. 2016. Application of data mining classification in employee performance prediction. *International Journal of Computer Applications*, 146, 28-35.

Lavrač, N., Bohanec, M., Pur, A., Cestnik, B., Debeljak, M. & Kobler, A. 2007. Data mining and visualization for decision support and modeling of public health-care resources. *Journal of Biomedical Informatics*, 40, 438-447.

Li, B., Akintoye, A., Edwards, P. J. & Hardcastle, C. 2005. Critical success factors for PPP/PFI projects in the UK construction industry. *Construction Management and Economics*, 23, 459-471.

Li, X., Zhu, Z. & Pan, X. 2010. Knowledge cultivating for intelligent decision making in small & middle businesses. *Procedia Computer Science*, 1, 2479-2488.

Marzouk, M. & Fayez, E. 2018. Public private partnership projects: concessionaire performance measurement. *Journal of Al Azhar University Engineering Sector*, 13, 466-480.

Mcquaid, R. W. & Scherrer, W. 2010. Changing reasons for public–private partnerships (PPPs). *Public Money and Management*, 30, 27-34.

www.SID.ir

Archive of SID

Mutula, S. M. & Van Brakel, P. 2006. E-readiness of SMEs in the ICT sector in Botswana with respect to information access. *The Electronic Library*, 24**,** 402-417.

Ogwueleka, T. C. & Ogwueleka, F. N. 2010. Data mining application in predicting *Cryptosporidium spp. oocysts and Giardia spp. cysts* concentrations in rivers. *Journal of Engineering Science and Technology*, 5**,** 342-349.

Qiao, L., Wang, S. Q., Tiong, R. L. & Chan, T. S. 2001. Framework for critical success factors of BOT projects in China. *The Journal of Structured Finance*, 7**,** 53-61.

Qiu, Y., Li, J., Huang, X. & Shi, H. 2018. A feasible data-driven mining system to optimize wastewater treatment process design and operation. *Water*, 10**,** 1342.

Ribeiro, D., Sanfins, A. & Belo, O. 2013. Wastewater treatment plant performance prediction with support vector machines. *Industrial Conference on Data Mining*, 2013. Springer, 99-111.

Rokach, L. 2010. Ensemble-based classifiers. *Artificial Intelligence Review*, 33**,** 1-39.

Sachs, T., Tiong, R. & Qing Wang, S. 2007. Analysis of political risks and opportunities in public private partnerships (PPP) in China and selected Asian countries: survey results. *Chinese Management Studies*, 1**,** 126-148.

Salman, A. F., Skibniewski, M. J. & Basha, I. 2007. BOT viability model for large-scale infrastructure projects. *Journal of Construction Engineering and Management*, 133**,** 50-63.

Sun, J., Wang, R., Wang, X., Yang, H. & Ping, J. 2014. Spatial cluster analysis of bursting pipes in water supply networks. *Procedia Engineering*, 70**,** 1610-1618.

Wen, Y. Y., Huang, W. M., Wu, J., Chen, Y. & Song, J. Q. 2013. Water consumption analysis system based on data mining. *Applied Mechanics and Materials*, 241-244**,** 1093-1097.

Yu, I., Kim, K., Jung, Y. & Chin, S. 2007. Comparable performance measurement system for construction companies. *Journal of Management in Engineering*, 23**,** 131-139.

Yuan, J., Skibniewski, M. J., Li, Q. & Zheng, L. 2009. Performance objectives selection model in public-private partnership projects based on the perspective of stakeholders. *Journal of Management in Engineering*, 26**,** 89-104.

Yuan, J., Wang, C., Skibniewski, M. J. & Li, Q. 2011. Developing key performance indicators for public-private partnership projects: questionnaire survey and analysis. *Journal of Management in Engineering*, 28**,** 252-264.

Zhang, X. 2005. Critical success factors for public–private partnerships in infrastructure development. *Journal of Construction Engineering and Management*, 131**,** 3-14.

www.SID.ir