

## Framework of Using Data Mining for Business Process Improvement

(Document Type: Research Paper)

**Mohammad Khanbabaei**

PhD Student, Department of Information Technology Management, Science and Research Branch, Islamic Azad University, Tehran, Iran, mohammadkhanbabaei@srbiau.ac.ir

**Farzad Movahedi Sobhani\***

Assistant professor, Department of Industrial Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran, farzadmovahedi@gmail.com

**Mahmood Alborzi**

Associate Professor, Department of Information Technology Management, Science and Research Branch, Islamic Azad University, Tehran, Iran, m.alborzi@srbiau.ac.ir

**Reza Radfar**

Associate Professor, Department of Technology Management, Science and Research Branch, Islamic Azad University, Tehran, Iran, r.radfar@srbiau.ac.ir

**Abstract:** Organizations should consider process improvement for promoting their performance. The main problem is the high volume of process dataset along with the wide variety of their features that leads to increase the complexity in process improvement methods. Previous methods cannot improve processes in a high volume of process dataset. Data mining approach can support improvement methods by identifying valuable hidden patterns in high volume of process dataset. In this paper a framework of using data mining techniques for extracting valuable hidden patterns in high volume of process dataset is developed for presenting the improvement suggestions. In order to evaluate the proposed framework, a real set of processes along with their features were gathered. After that, using classification, clustering, and features selection algorithms, valuable patterns in high volume of process dataset were identified. After evaluating these patterns, improvement suggestions were recommended by the identified patterns. The results showed that identified patterns can support processes improvement activities with recommending the improvement suggestions.

**Keywords:** Process Improvement, Data Mining, Business Processes.

**Introduction:** Processes are one of the most important resources in organizations. Improving processes can lead to enhance the organizational performance. Several methodologies are presented for process improvement. However, they did not consider the problem of high volume of process dataset along with their features in organizations. There are many processes in organizations that lead to increase the complex interactions between processes and high dimensionality problem (Jeong et al., 2008). In addition, there is a single view to processes in the improvement actions (Houy et al., 2011). Huang et al. (2012) stated that there is a low attention to the internal aspects of the processes. In this situation, data mining techniques can identify and discover valuable patterns hidden in the large number of processes in the organizations. These patterns can be utilized for recommending the improvement suggestion for enhancing the performance of the processes. Recently, there is a link between data mining and process improvement. A few studies considered the interaction between data mining and process improvement. However, these studies did not consider a real large number of processes in their computations. In addition, they did not have a comprehensive and practical view to the application of data mining for process improvement. This paper presents a framework for using data mining techniques for identifying and extracting valuable and suitable patterns hidden in the large number of processes. These patterns can be employed for recommending the process improvement suggestions.

**Materials and Methods:** This paper employs three data mining techniques including clustering, classification, and feature selection techniques for extracting valuable patterns hidden in the large number of processes. CRISP-DM (cross industry standard process for data mining) standard is used to implement data mining activities. In the classification technique, C5 decision tree algorithm is employed to classify processes. In clustering, K-means algorithm is applied to segment processes in several clusters. In feature selection technique, the most important process features are selected for process improvement. In the proposed framework, at first, all processes are gathered from a variety of the resources in the organization. After that, process features are identified from the literature and they are defined based on an interaction between data miner and process improvement expert. In the following, a process dataset is provided for using data mining techniques. For this

---

\* Corresponding author

Copyright © 2019, University of Isfahan. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits others to download this work and share it with others as long as they credit it, but they cannot change it in any way or use it commercially.

dataset, a variety of data preparation and preprocessing methods are employed for achieving better results of implementing data mining techniques. After that, three classification, clustering, and feature selection data mining techniques are applied for extracting valuable patterns hidden in the large number of processes. In the classification techniques, C5 decision tree algorithm employs a target features for classifying processes. Cross validation method is applied to train and test the constructed decision tree. The output is several if-then rules for classifying processes based on the target feature. In clustering, K-means algorithm segments processes in several clusters. Processes in a cluster have a similar behavior and they are dissimilar to the processes in the other clusters. Euclidean distance function is used for calculating the distance between processes. The output is the cluster profiling to describe the behavior of processes in each cluster. In the feature selection, the most important features are selected based on a target process feature. These features are more correlated to the target process feature. In addition, they are more important to consider for process improvement purposes. The output is a variety of more important process features so that they can be considered for recommending the process improvement suggestions. After extracting valuable patterns hidden in the large number of processes, the accuracy and quality and these patterns are evaluated by an interaction between data miner, process owner, and process improvement expert. Evaluated patterns are considered for recommending the improvement suggestions. These suggestions must be aligned with the process improvement concepts in the organization. Processes are improved based on these suggestions. In last, the performance of the improved processes is evaluated. The proposed framework is based on an iterative and continuous method for using data mining in process improvement.

**Results and Discussion:** The proposed framework was evaluated based on a real process dataset including 1318 processes and 80 process features. Several preprocessing methods were employed to prepare dataset. Three classification, clustering, and feature selection techniques were applied to extract valuable patterns hidden in the large number of processes. Using C5 classification algorithm, processes were classified by the if-then rules based on a target process feature. In the proposed model, as an example, key processes were classified using C5 decision tree. The classification accuracy in the test dataset was set to 92.31%. The output was several if-then rules to classify key processes. These rules can identify key processes. In addition, processes features applied to construct the if-then rules can be employed for recommending the improvement suggestions for key processes. In clustering, using K-means algorithm, processes were segmented into 10 clusters. Some features were considered to cluster processes. The output was cluster profiling to describe the behavior of processes based on the selected features. In this paper, for example, the behaviors of processes in cluster 1 were described through the selected features. Several suggestions were recommended by the selected features for improving processes in this cluster. In the feature selection technique, for example, “cost of process” was considered as the target process feature. The output was the selection of 10 more important process features that they were more related to the target process features. These features can describe the cost of processes better than the other features. The improvement suggestions for reducing cost of processes were recommended based on these selected features.

**Conclusion:** This paper presented a framework of using data mining techniques for identifying valuable patterns hidden in the large number of processes for the process improvement purposes. A real process dataset was employed to evaluate the applicability and effectiveness of the proposed framework. In the proposed framework, classification, clustering, and feature selection data mining techniques were applied to extract valuable patterns hidden in high volume of process dataset. Process improvement methodologies cannot recommend the improvement suggestions in a rapid and accurate method, when there are many processes along with a variety of the process features. These methodologies are restricted to recommend a limited number of the process improvement suggestions. In the other direction, there are few studies on the application of data mining for the process improvement that they include some weaknesses. The proposed framework employed a lot of real processes in the data mining techniques to discover valuable patterns for process improvement. In addition, a variety of the process features were extracted from the literature to describe the behavior of the processes. In last, a wide variety of the several suggestions were recommended for process improvement based on the extracted patterns hidden in the large number of processes. The organizations can utilize the proposed framework for improving their processes. In addition, this framework can help organizations with the large number of processes for employing the process improvement methodologies in a productive manner. Future studies can apply the other data mining techniques for the proposed framework. In addition, the proposed framework can be developed for knowledge-intensive processes. The proposed framework can be integrated with the knowledge management methodologies for improving knowledge-based processes. In last, the proposed framework can be enhanced as a decision support system for the process improvement.

## References

- Houy, C., Fettke, P., Loos, P., van der Aalst, W.M.P., Krogstie, J. (2011). Business Process Management in the Large. *Business and Information Systems Engineering*, 3, 385–388.
- Huang, Z., Lu, X., Duan, H. (2012). Resource behavior measure and application in business process management. *Expert Systems with Applications*, 39 (7), 6458–6468.
- Jeong, H., Song, S., Shin, S., Rae Cho, B. (2008). Integrating data mining to a process design using the robust bayesian approach. *International Journal of Reliability, Quality and Safety Engineering*, 15 (05), 441–464.

مدیریت تولید و عملیات، دوره ۱۰، پیاپی ۱۸، شماره ۱، بهار و تابستان ۱۳۹۸

دریافت: ۱۳۹۵/۱۱/۱۶ پذیرش: ۱۳۹۷/۰۹/۱۷

صص: ۲۵-۴۵ (نوع مقاله: پژوهشی)

## چارچوب به کارگیری داده‌کاوی برای بهبود فرآیندهای کسب‌وکار

محمد خان‌بابایی<sup>۱</sup>، فرزاد موحدی سبحانی<sup>۲\*</sup>، محمود البرزی<sup>۳</sup>، رضا رادفر<sup>۴</sup>

۱- دانشجوی دکتری گروه مدیریت فناوری اطلاعات، دانشگاه آزاد اسلامی، واحد علوم و تحقیقات، تهران، ایران،

mohammadkhanbabaei@srbiau.ac.ir

۲- استادیار گروه مهندسی صنایع، دانشگاه آزاد اسلامی، واحد علوم و تحقیقات، تهران، ایران، fmovahedi@iau.ac.ir

۳- دانشیار گروه مدیریت فناوری اطلاعات، دانشگاه آزاد اسلامی، واحد علوم و تحقیقات، تهران، ایران، m.alborzi@srbiau.ac.ir

۴- دانشیار گروه مدیریت فناوری اطلاعات، دانشگاه آزاد اسلامی، واحد علوم و تحقیقات، تهران، ایران، r.radfar@srbiau.ac.ir

**چکیده:** سازمان‌ها برای ارتقای عملکرد خود باید به بهبود فرآیندها توجه داشته باشند. مسئله اصلی حجم زیاد فرآیندها همراه با تنوع وسیع ویژگی‌های آنها است که باعث افزایش پیچیدگی در روش‌های بهبود فرآیندها می‌شود. روش‌های قبلی به بهبود فرآیندها در حجم زیاد فرآیندها قادر نیستند. رویکرد داده‌کاوی با شناسایی الگوهای ارزشمند پنهان در حجم زیاد فرآیندها از روش‌های بهبود پشتیبانی می‌کند. در این مقاله چارچوبی برای به کارگیری روش‌های داده‌کاوی برای استخراج الگوهای ارزشمند پنهان در حجم زیاد فرآیندها با هدف ارائه پیشنهادهای بهبود توسعه داده شده است. برای ارزیابی چارچوب پیشنهادی، مجموعه‌ای واقعی از فرآیندها به همراه ویژگی‌های آنها جمع‌آوری شده است. سپس با الگوریتم‌های طبقه‌بندی، خوشه‌بندی و انتخاب ویژگی‌ها، الگوهایی با ارزش در حجم زیاد فرآیندها شناسایی شدند. بعد از ارزیابی این الگوها، پیشنهادهای بهبود از طریق الگوهای شناسایی شده پیشنهاد شده است. نتایج نشان می‌دهد الگوهای شناسایی شده قادر هستند با ارائه پیشنهادهای بهبود اقدامات بهبود فرآیندها را پشتیبانی کنند.

**واژه‌های کلیدی:** بهبود فرآیندها، داده‌کاوی، فرآیندهای کسب‌وکار

## مقدمه

فرآیندها یکی از دارایی‌های راهبردی‌اند که نقش مهمی در سازمان دارند. هارینگتون<sup>۱</sup> (۱۹۹۱) بیان کرده است فرآیند، فعالیت یا گروهی از فعالیت‌ها است که ورودی (مثل اطلاعات، مستندات، منابع مالی، نیروی انسانی) را دریافت و برای افزایش ارزش فرآیند آن را پردازش می‌کند تا خروجی (مثل تصمیم‌گیری، دانش، خدمات، محصول، مستندات) مدنظر را تأمین کند؛ برای مثال ورودی «فرآیند تست» شامل مستندات و نیروی انسانی متخصص برای انجام تست است. خروجی این فرآیند شامل نتایج تست و تصمیم‌گیری درباره آن است.

سازمان‌ها برای دستیابی به اهداف خود نیازمند بهبود فرآیندها هستند. دالماریس<sup>۲</sup> و همکاران (۲۰۰۷) معتقدند بهبود فرآیندها از سال ۱۹۹۰ با کار اندیشمندانی چون داوِنپورت، هَمِر و چَمپی<sup>۳</sup> توسعه یافته است. هَمِر (۲۰۰۲) بیان کرد بهبود فرآیندها رویکردی ساخت‌یافته برای ارتقای عملکرد و متمرکز بر طراحی منظم و اجرای دقیق فرآیندها است.

چارچوب‌های مختلفی برای بهبود فرآیندها ارائه شده است. چارچوب هرم روندهای فرآیندهای کسب‌وکار را هارمون<sup>۴</sup> (۲۰۰۷) ارائه کرده است. همچنین تقسیم‌بندی‌های مختلفی از فرآیندها انجام شده است. برن<sup>۵</sup> (۲۰۰۸) فرآیندها را به دو دسته کلیدی و پشتیبانی تقسیم کرده است. رنجبرفرد<sup>۶</sup> و همکاران (۲۰۱۳) فرآیندها را به سه دسته تقسیم کرده‌اند؛ این سه دسته عبارتند از معمولی، تکرارشونده و غیرتکرارشونده.

در این چارچوب‌ها و تقسیم‌بندی‌ها، مسئله حجم بسیار زیاد فرآیندها و ویژگی‌های آنها که باعث ایجاد تعاملات پیچیده بین فرآیندها می‌شود، در نظر گرفته نشده است. در یک سازمان مجموعه بزرگی از فرآیندها به همراه داده‌ها و ویژگی‌های آنها وجود دارد. جونگ<sup>۷</sup> و همکاران (۲۰۰۷) بیان کردند داده‌های فرآیند و نتایج درک‌شده یا مشاهده‌شده از آن، دانش ارزشمندی برای سازمان‌اند. همچنین چِن و وِنگ<sup>۸</sup> (۱۹۹۹) بیان کردند فرآیندها در سازمان‌ها دارای ویژگی‌هایی مانند حجم زیاد، مسئله ابعاد بالای فرآیند، ماهیت پویای ویژگی‌های فرآیندها، تعاملات پیچیده بین ویژگی‌های فرآیندها و وجود مقیاس‌های اندازه‌گیری متعدد فرآیندها هستند.

جونگ<sup>۹</sup> و همکاران (۲۰۰۸) بیان کردند با افزایش تعداد فرآیندها و ویژگی‌های آنها مسئله ابعاد زیاد رخ می‌دهد. همچنین هوی<sup>۱۰</sup> و همکاران (۲۰۱۱) معتقد بودند در گذشته، مدیریت فرآیندها صرفاً نگاه تکی و مجزا به فرآیندها داشتند؛ اما مدیریت فرآیندها در آینده متمرکز بر مجموعه‌های بزرگ از فرآیندها است. لیمِتس<sup>۱۱</sup> و همکاران (۲۰۱۲) اعتقاد داشتند بهبود فرآیندها بدون در نظر گرفتن تأثیر آنها بر هم انجام می‌شود. از طرف دیگر به جنبه‌های درونی فرآیند (ویژگی‌های فرآیند) توجه کمی شده است (هوآنگ<sup>۱۲</sup> و همکاران، ۲۰۱۲).

همچنین ووکسیک<sup>۱۳</sup> و همکاران (۲۰۱۳) بیان کردند در گذشته ارتباطی بین داده و فرآیند در رویکردهای هوشمندی کسب‌وکار وجود نداشته است. آنها معتقدند اجرای مجزای هوشمندی کسب‌وکار و مدیریت فرآیندها باعث بهبود عملکرد و کاهش هزینه‌ها در سازمان نمی‌شود. دلگادو<sup>۱۴</sup> و همکاران (۲۰۱۴) معتقد بودند نگرشی یکپارچه و تصویر کامل برای تحلیل فرآیندها براساس اطلاعات آنها وجود ندارد. همچنین در بیشتر روش‌شناسی‌های بهبود فرآیندها، هوشمندی وجود ندارد.

از طرف دیگر، داده‌کاوی در حجم زیاد فرآیندها باعث شناسایی و کشف الگوهای پنهان می‌شود. این الگوها به سازمان در بهبود فرآیندها کمک می‌کنند. ژنگوا و لیمئی<sup>۱۵</sup> (۲۰۰۸) معتقدند داده‌کاوی می‌تواند ویژگی‌های فرآیندها

را شناسایی و از تحلیل عوامل کلیدی موفقیت، شناسایی فرآیندهای کلیدی، بهبود جریان اطلاعات و بازخوردهای آنها پشتیبانی کند. فلُرونسو و اُگوندا<sup>۱۶</sup> (۲۰۰۵) و تن<sup>۱۷</sup> و همکاران (۲۰۰۶) بیان کردند استخراج الگوها در مجموعه فرآیندها باعث هدایت مدیران در بهبود فرآیندها می‌شود. تیواری<sup>۱۸</sup>، و همکاران (۲۰۰۸) معتقد بودند داده‌کاوی به تحلیل رفتار فرآیندها، شناسایی دوباره‌کاری، وجود اختلال، حلقه‌ها، توالی و هم‌زمانی در فرآیندها و تجسم‌سازی آنها می‌پردازد. وِگنر و راپینگ<sup>۱۹</sup> (۲۰۱۰) بیان کردند داده‌کاوی در مواردی مانند حذف فعالیت‌های اضافی، حذف فعالیت‌های موازی، شناسایی فعالیت‌های تکراری در سازمان، کاهش تعداد افراد یا واحدهای درگیر در یک فرآیند و تحلیل فرآیندها در یک الگوی مشخص، به بهبود فرآیندها کمک می‌کند.

امروزه بهبود فرآیندها با موضوع داده‌کاوی پیوند خورده است. باتوجه‌به حجم زیاد فرآیندها و ویژگی‌های آنها، پژوهش‌های اندکی موضوع بهبود فرآیندها از طریق داده‌کاوی را بررسی کرده‌اند. این پژوهش‌ها نیز ارتباط بین الگوریتم‌های داده‌کاوی و بهبود فرآیندها را بررسی نکرده‌اند؛ بلکه فقط از یک روش داده‌کاوی برای انجام کاری جزئی روی فرآیندها استفاده کرده‌اند. نکته درخور توجه، سرعت داده‌کاوی در پشتیبانی از بهبود فرآیندها بدون درگیر بودن همه سازمان نسبت به سایر روش‌های بهبود فرآیندها است. از طریق داده‌کاوی الگوهای پنهان در حجم زیاد فرآیندها به سرعت شناسایی می‌شود و براساس آنها استراتژی‌های بهبود فرآیندها انتخاب و اجرا می‌شود. مزیت اصلی دیگر شناخت ارتباط بین فرآیندها است. داده‌کاوی الگوهای پنهان در ارتباط بین فرآیندها را برای بهبود آنها کشف می‌کند.

در این مقاله چارچوبی برای به‌کارگیری روش‌های داده‌کاوی در شناسایی و استخراج الگوهای مفید پنهان در مجموعه فرآیندها کشف شده است؛ به‌گونه‌ای که با این الگوها پیشنهادهایی برای بهبود فرآیندها ارائه می‌شود. هدف اصلی کشف الگوهای ارزشمند در مجموعه فرآیندها، ارائه پیشنهاد برای بهبود آنها است. برای دستیابی به این هدف باید در ابتدا الگوهای رفتاری پنهان در حجم زیاد فرآیندها کشف شود. سپس، پیشنهادهایی برای بهبود فرآیندها براساس الگوهای رفتاری کشف‌شده ارائه شود؛ بنابراین دو سوال اصلی در این مقاله مطرح می‌شود: ۱. چگونه الگوهای مفید در حجم زیاد فرآیندها شناسایی شود؟ ۲. چگونه با الگوهای شناخته‌شده، پیشنهادهایی برای بهبود فرآیند ارائه می‌شود؟

## مفاهیم و پیشینه پژوهش

در این بخش مفاهیم به‌کاررفته در مقاله ارائه می‌شود.

## فرآیند و بهبود فرآیند

فرآیند مفهومی مشخص شامل یک‌سری پیش‌فرض‌ها، نتایج، محتوا، اقدامات و دلایل است (گومز پرز<sup>۲۰</sup> و همکاران، ۲۰۱۰). همچنین بُرگو و باربا<sup>۲۱</sup> (۲۰۱۴) بیان کرده‌اند فرآیند اقدامات انجام‌شده در محیطی فنی و سازمانی برای تحقق اهداف کسب‌وکار است. دمیچ و دمیچ<sup>۲۲</sup> (۲۰۱۴) بیان کرده‌اند بهبود فرآیند متمرکز بر بهبود کارکرد فرآیندهای جاری با جستجوی روش‌هایی برای افزایش عملکرد و کیفیت و کاهش هزینه‌ها است. برخی از این روش‌ها عبارتند از حذف بوروکراسی؛ تحلیل ارزش افزوده؛ حذف دوباره‌کاری؛ ساده‌کردن روش‌ها؛ کاهش زمان؛ اصلاح خطاها؛ به‌روزرسانی فرآیند؛ استانداردسازی؛ مشارکت تأمین‌کنندگان؛ به‌کارگیری فناوری اطلاعات در اجرای فرآیند.

## داده‌کاوی

داده‌کاوی، فرآیند انتخاب، کشف و مدل‌سازی برای یافتن الگوهای پنهان در حجم زیاد داده است (کوه و لو<sup>۳۳</sup>، ۲۰۰۴). این الگوها صریح، مفید و بالقوه‌اند (لی و ساوو<sup>۲۴</sup>، ۲۰۰۱). لارنس (۲۰۰۵) فرآیند کریسپ<sup>۲۵</sup> برای اجرای داده‌کاوی را تعریف کرده است؛ این تعریف شامل ۱. فهم کسب‌وکار؛ ۲. فهم داده؛ ۳. آماده‌سازی داده؛ ۴. مدل‌سازی؛ ۵. ارزیابی و ۶. به‌کاربری است.

روش‌های متعددی در داده‌کاوی وجود دارند که در این مقاله از سه روش طبقه‌بندی، خوشه‌بندی و انتخاب ویژگی برای یافتن الگوهای ارزشمند در حجم زیاد فرآیندها برای بهبود آنها استفاده می‌شود.

در طبقه‌بندی، الگوریتم درخت تصمیم‌گیری، رکوردهای مجموعه‌ای داده را براساس یک متغیر هدف طبقه‌بندی می‌کند. خروجی این الگوریتم، مجموعه‌ای از قوانین اگر-آنگاه است. این الگوریتم از فرآیند استنتاج بالا به پایین براساس روش بخش‌بندی بازگشتی بهره می‌برد (دی هیگر<sup>۲۶</sup> و همکاران، ۲۰۰۳). درخت تصمیم‌گیری شامل چندین گره، شاخه و برگ است. هر گره معرف یک متغیر است. شاخه‌ها مجموعه داده‌ها را به مجموعه‌های کوچک‌تر تقسیم می‌کنند. برگ‌ها در انتهای شاخه‌ها، رکوردها را براساس متغیر هدف طبقه‌بندی می‌کنند (دی هیگر و همکاران، ۲۰۰۳).

روش خوشه‌بندی رکوردهای مجموعه داده را به چند خوشه بخش‌بندی می‌کند. بدین منظور فاصله بین رکوردها محاسبه می‌شود و رکوردهایی که فاصله بین آنها کم است در یک خوشه واقع می‌شوند. این رویه تا محقق شدن هدف در الگوریتم خوشه‌بندی ادامه می‌یابد (لارنس، ۲۰۰۵).

روش انتخاب ویژگی، ویژگی‌های (متغیرهای) مهم را براساس یک ویژگی هدف انتخاب می‌کند. الگوریتم انتخاب ویژگی دارای سه مؤلفه است (سالایا<sup>۲۷</sup> و همکاران، ۲۰۰۷)؛ این مؤلفه‌ها عبارتند از: ۱. مقیاس ارزیابی ویژگی؛ ۲. روش جستجو؛ ۳. معیار توقف.

## پیشینه پژوهش

برخی پژوهش‌ها کاربرد داده‌کاوی در بهبود فرآیندها را بررسی کرده‌اند. چن و ونگ (۱۹۹۹) از داده‌کاوی در تحلیل داده‌های عملیاتی فرآیند استفاده کرده‌اند. وگنر و راپینگ (۲۰۱۰) مدلی برای یکپارچگی داده‌کاوی و فرآیندها ارائه کرده‌اند. نیسن<sup>۲۸</sup> (۱۹۹۹) با روش داده‌کاوی سیستمی مبتنی بر دانش را برای تشخیص مسائل فرآیندها با هدف بهبود آنها طراحی کرده است. فلرونسو و اگوندا (۲۰۰۵) از داده‌کاوی برای پشتیبانی طراحی مجدد فرآیند با استخراج الگو استفاده کرده است. ژنگوا و لیمئی (۲۰۰۸) کاربرد داده‌کاوی در مهندسی مجدد فرآیندها را برای تحلیل داده از طریق شناسایی فرآیندهای کلیدی، تحلیل عوامل بحرانی موفقیت، بهبود جریان اطلاعات و بازخورد فرآیندها بررسی کرده است. کارکو<sup>۲۹</sup> و همکاران (۲۰۰۸) مفهوم فناوری‌های هوشمند فرآیند همچون داده‌کاوی را ارائه کرده‌اند.

گریگوری<sup>۳۰</sup> و همکاران (۲۰۰۴) کاربرد هوشمندی کسب‌وکار در فرآیندها را بررسی کرده‌اند. میو و جرج<sup>۳۱</sup> (۲۰۱۲) بیان کرده‌اند می‌توان از داده‌کاوی برای یافتن دانش پنهان استفاده کرد و از آن در مهندسی مجدد فرآیندها بهره برد. قنادباشی<sup>۳۲</sup> و همکاران (۲۰۱۳) مدل یکپارچگی داده‌کاوی و مهندسی مجدد فرآیندها را ارائه کرده‌اند. در

آنجا تبیین شده است که داده‌کاوی چگونه می‌تواند در هر فاز مهندسی مجدد فرایندها، الگوهای مناسب در مجموعه فرایندها را شناسایی کند. گروگر<sup>۳۳</sup> و همکاران (۲۰۱۴) برای بهبود فرایندها از طریق سیستم‌های پیشنهاددهنده درخت تصمیم‌گیری، تحلیل تجویزی را ارائه کرده‌اند. پیک<sup>۳۴</sup> و همکاران (۲۰۱۴) رویکردی برای اجرای داده‌کاوی در فرایندها ارائه کرده‌اند که شامل فازهای ارزیابی، بازطراحی، مدل‌سازی و قابلیت اجرایی شدن نتایج داده‌کاوی است.

راپنیک و جکلک<sup>۳۵</sup> (۲۰۰۹) کاربرد داده‌کاوی در فرایندها را به صورت کلی ارائه کرده‌اند و از استاندارد کریسپ بهره برده‌اند. وگنر و رایپنگ (۲۰۱۰) توصیفی تئوریک از ارتباط بین داده‌کاوی و فرایندهای مبتنی بر استاندارد کریسپ ارائه داده‌اند؛ اما ابعاد فنی، خروجی‌ها و کاربرد داده‌کاوی در بهبود فرایندها را بیان نکرده‌اند. سهیل و هاناپال دورای دومینیک<sup>۳۶</sup> (۲۰۱۲) مدلی برای کاهش شکاف بین هوشمندی فرآیند و بهبود فرآیند ارائه کرده‌اند. قاتاس<sup>۳۷</sup> و همکاران (۲۰۱۴) با استفاده از روش‌های داده‌کاوی رویکردی نیمه‌خودکار برای بهبود عملکرد براساس تصمیمات در خصوص فرایندهای قبلی ارائه کرده‌اند.

### چارچوب نظری و وجه تمایز مطالعه

شکل ۱ چارچوب نظری مستخرج از پیشینه پژوهش را نشان می‌دهد. این چارچوب ارتباط بین کارکردهای داده‌کاوی و بهبود فرایندها را ارائه می‌کند. کارکردها، دارای تعدادی ویژگی و رفتارند. ویژگی‌ها نشان‌دهنده مشخصه‌های کارکردها هستند. رفتارها مجموعه اعمالی است که هر یک از کارکردها در مواجهه با یکدیگر انجام می‌دهند. نوآوری این مقاله، ارائه چارچوبی برای به‌کارگیری داده‌کاوی برای بهبود فرایندها است. در این چارچوب از داده‌کاوی برای شناسایی الگوهای پنهان در حجم زیاد فرایندها استفاده می‌شود؛ با استفاده از این الگوها، پیشنهادهای بهبود برای فرایندها ارائه می‌شود.



شکل ۱- چارچوب نظری ارتباط بین کارکردهای داده‌کاوی و بهبود فرایندها

وجه تمایز این مطالعه از دو جنبه بررسی شدنی است؛ نخست روش‌های بهبود فرآیندها، مسئله حجم زیاد فرآیندها و ویژگی‌های زیاد آنها را در نظر نمی‌گیرند. این مقاله چارچوبی برای شناسایی الگوهای پنهان در حجم زیاد فرآیندها با داده‌کاوی ارائه می‌دهد. به کمک این الگوها، پیشنهادهاى مختلف برای بهبود فرآیندها ارائه می‌شود؛ بنابراین این مطالعه براساس چارچوبی جدید از روش‌شناسی‌های بهبود در حجم زیاد فرآیندها پشتیبانی می‌کند. دوم، روش‌شناسی‌های بهبود، فرآیندها را فقط با یک یا چند ویژگی فرآیندی تحلیل می‌کند؛ اما چارچوب پیشنهادی تعداد زیادی ویژگی (برگرفته از مفاهیم بهبود فرآیندها) را لحاظ می‌کند.

نکته درخور توجه این است که فرآیندکاوی یکی از رویکردهای مشهور در زمینه بهبود فرآیندها است. فرآیندکاوی به کشف، پایش و بهبود فرآیندها از طریق دانش مستخرج از داده‌های ثبت واقعه می‌پردازد (کلاس و پولس<sup>۳۸</sup>، ۲۰۱۴). اگرچه مطالعات متعددی بر فرآیندکاوی متمرکز هستند، این مقاله قصد ندارد جایگزینی برای فرآیندکاوی ارائه کند. ایده مقاله به جای فرآیندکاوی، متمرکز بر بهبود فرآیند است. براساس جدول ۱، چارچوب پیشنهادی در چندین موضوع با فرآیندکاوی متفاوت است.

جدول ۱- تفاوت چارچوب پیشنهادی با فرآیندکاوی

مطالعات فرآیندکاوی	نویسنده (سال)	چارچوب پیشنهادی
فقدان به‌کارگیری وسیع مفاهیم بهبود فرآیندها	دوماس <sup>۳۹</sup> و همکاران (۲۰۱۳)	ارتباط فرآیندها با مفاهیم بهبود فرآیندها
استفاده خفیه کم از ویژگی‌های فرآیندها	سونگ <sup>۴۰</sup> و ون در آلت (۲۰۰۸)	به‌کارگیری تنوع زیادی از ویژگی‌های فرآیندها
ارائه تعداد محدودی پیشنهادهاى بهبود فرآیند	دوماس و همکاران (۲۰۱۳)	ارائه تعداد زیاد و متنوع پیشنهادهاى بهبود فرآیند
استخراج الگوهای سخت از اطلاعات ثبت وقایع	کاستیلانوس <sup>۴۱</sup> و همکاران (۲۰۰۴)؛ ینگ <sup>۴۲</sup> و همکاران (۲۰۱۴)	استخراج الگوهای منعطف از تعداد بسیار زیاد فرآیندها
استفاده کم از نظرات کارکنان و مدیران سازمان	کاستیلانوس و همکاران (۲۰۰۴)؛ ینگ و همکاران (۲۰۱۴)	دریافت نظرات کارشناسان در تحلیل فرآیندها، تکمیل اطلاعات فرآیندها، به‌کارگیری الگوهای داده‌کاوی و ارائه پیشنهادهاى بهبود
تمرکز بر جنبه درونی فرآیندها	سونگ و ون در آلت (۲۰۰۸)	توجه زیاد به ابعاد بیرونی فرآیندها همچون ابعاد سازمانی و اهداف فرآیندها و کسب‌وکار و همچنین ویژگی‌های فرآیندها
فهم کم از فرآیندهای کسب‌وکار	دوماس و همکاران (۲۰۱۳)	فهم وسیع از فرآیندها با به‌کارگیری ویژگی‌های متنوع فرآیندها
فهم کم خروجی‌های فرآیندکاوی برای سازمان	دوماس و همکاران (۲۰۱۳)	ارائه الگوها و نتایج قابل فهم بهبود فرآیند با داده‌کاوی

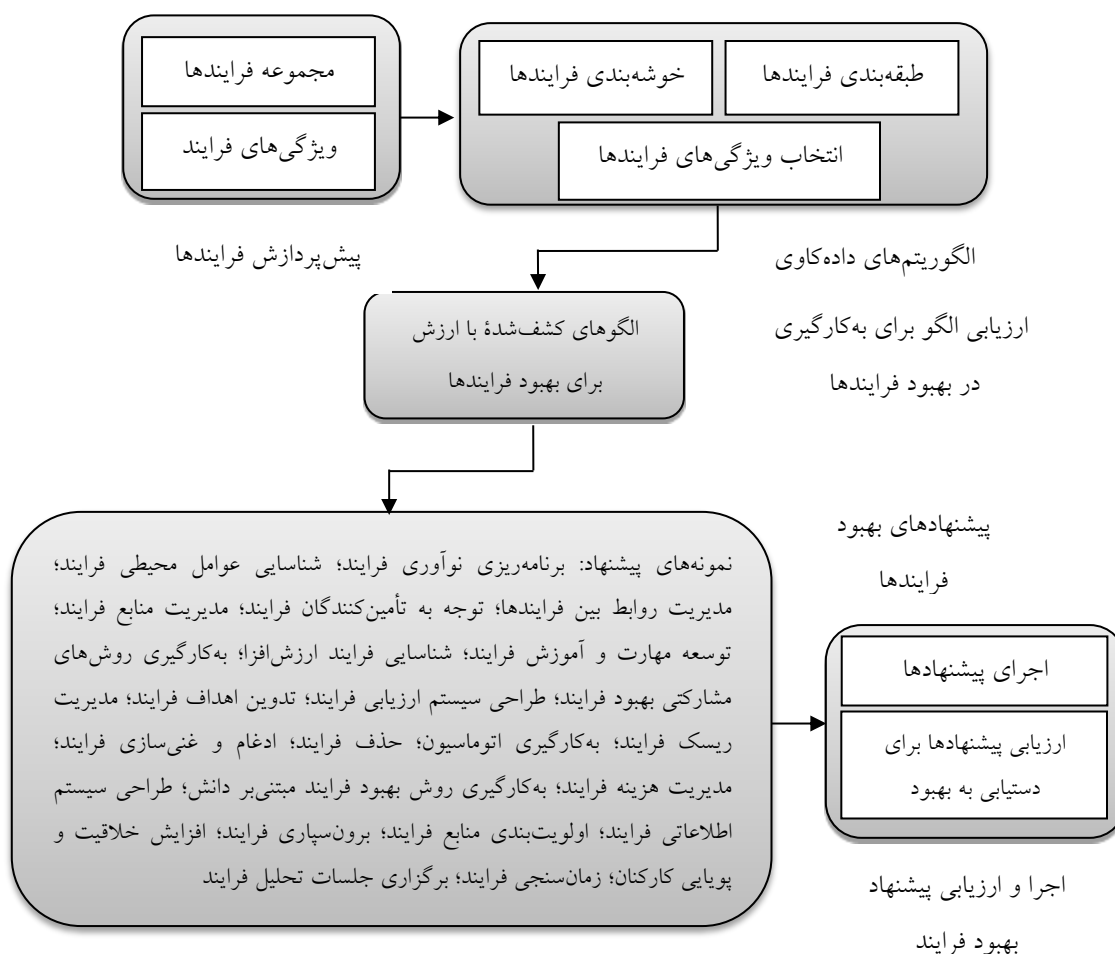
## روش‌شناسی

روش‌شناسی موضوع این مقاله از نظر هدف، توصیفی و از نظر نتایج، کاربردی است. از نظر فرآیند اجرا، مدل‌سازی مبتنی بر داده‌کاوی است. مراحل روش‌شناسی عبارتند از ۱. شناسایی مجموعه فرآیندهای سازمان؛ ۲. شناسایی ویژگی‌های فرآیندها؛ ۳. تهیه بانک اطلاعاتی مشتمل بر مجموعه فرآیندها، ویژگی‌ها و مقادیر آنها؛ ۴. عملیات آماده‌سازی و پیش‌پردازش برای تهیه مجموعه با کیفیت از بانک اطلاعاتی فرآیندها؛ ۵. اجرای روش‌های



داده‌کاوی طبقه‌بندی، خوشه‌بندی و انتخاب ویژگی برای یافتن الگوهای مفید پنهان در فرایندها؛ ۶. ارزیابی دقت و کیفیت خروجی‌های داده‌کاوی (الگوهای کشف‌شده)؛ ۷. استنتاج الگوهای کشف‌شده برای ارائه پیشنهادها به‌بهبود فرایندها (به‌وسیله خبرگان)؛ ۸. ارائه پیشنهادها به‌بهبود از طریق الگوهای کشف‌شده.

شکل ۲ چارچوبی برای به‌کارگیری داده‌کاوی در بهبود فرایندها ارائه می‌کند. در ابتدا فرایندها شناسایی می‌شوند. مرحله بعد تعیین ویژگی‌های فرایندها و تعریف عملیاتی آنها است که از متخصص فرآیند و کارشناس فرایندها استفاده می‌شود. ویژگی‌های فرایندها برگرفته از مفاهیم بهبود فرایندها است. در ادامه بانک اطلاعاتی فرایندها به‌همراه ویژگی‌های آنها ایجاد می‌شود. مقادیر ویژگی‌ها با تعاریف عملیاتی آنها تعیین می‌شود. سپس عملیات آماده‌سازی و پیش‌پردازش فرایندها صورت می‌گیرد. با استفاده از الگوریتم‌های طبقه‌بندی، خوشه‌بندی و انتخاب ویژگی تنوعی از الگوهای پنهان در حجم زیاد فرایندها استخراج می‌شود. این الگوها در تحلیل رفتار فرایندها و ارائه پیشنهادها به‌بهبود به کار می‌روند.



شکل ۲- چارچوب پیشنهادی به‌کارگیری داده‌کاوی در بهبود فرایندها

الگوریتم‌های داده‌کاوی به‌کاررفته در چارچوب پیشنهادی به‌شرح زیر است:  
 ۱. الگوریتم طبقه‌بندی، مدل طبقه‌بندی فرایندها را براساس یک ویژگی هدف می‌سازد. در اینجا از الگوریتم طبقه‌بندی درختان تصمیم‌گیری استفاده می‌شود. این الگوریتم، کاربرپسند است و می‌تواند الگوهای ساده با تفسیر آسان و جالب در مجموعه داده تولید کند.

۲. الگوریتم خوشه‌بندی بدون در نظر گرفتن یک ویژگی هدف، بخش‌بندی فرایندها در خوشه‌های مختلف را انجام می‌دهد؛ به گونه‌ای که فرایندها در یک خوشه دارای بیشترین شباهت به هم هستند و بیشترین تفاوت بین آنها در خوشه‌های مختلف وجود دارد. برای خوشه‌بندی از الگوریتم کامیانگین استفاده می‌شود؛ این الگوریتم برای بخش‌بندی فرایندها محبوب و آسان است. این الگوریتم می‌تواند الگوهای با پیچیدگی کمتر و تفسیر ساده‌تر نسبت به سایر الگوریتم‌ها ایجاد کند.

۳. الگوریتم انتخاب ویژگی، ویژگی‌هایی را انتخاب می‌کند که نسبت به سایر ویژگی‌ها، همبستگی بیشتری با ویژگی هدف دارند و از اهمیت بیشتری برای بهبود فرایندها برخوردار هستند.

باتوجه به تنوع زیاد الگوریتم‌های داده‌کاوی، هدف اصلی به‌کارگیری الگوریتم‌های مختلف داده‌کاوی و ارزیابی آنها در یک نمونه مجموعه داده نیست. این مقاله متمرکز بر ارائه چارچوب به‌کارگیری داده‌کاوی در بهبود فرایندها است؛ بنابراین در این مقاله از سه الگوریتم طبقه‌بندی درخت تصمیم‌گیری سی‌پنج، خوشه‌بندی کامیانگین و الگوریتم انتخاب ویژگی مبتنی بر شاخص همبستگی ویژگی‌ها با ویژگی هدف در چارچوب پیشنهادی استفاده می‌شود. البته در اجرای چارچوب پیشنهادی می‌توان انواع مختلف الگوریتم‌های داده‌کاوی را به کار برد. باتوجه به اینکه یکی از مشکلات متخصصین بهبود فرآیند، عدم آشنایی آنها با الگوریتم‌های داده‌کاوی است، سعی می‌شود در اجرای چارچوب پیشنهادی از الگوریتم‌های ساده و کاربرپسند داده‌کاوی استفاده شود.

با اجرای داده‌کاوی، الگوهای مستخرج شناسایی می‌شوند. سپس با استفاده از روش‌های ارزیابی به شرح زیر، میزان دقت و کیفیت آنها تعیین می‌شود: ۱. نظرات خبرگان شامل متخصص بهبود، مالک و مدیر فرآیند و داده‌کاو برای ارزیابی مطابقت الگوها با مفاهیم بهبود فرایندها؛ ۲. بررسی پیچیدگی و قابلیت تفسیر الگوها به وسیله خبرگان؛ ۳. روش اعتبارسنجی متقاطع با عدد ۱۰ برای تقسیم مجموعه فرایندها به دو مجموعه آموزش و آزمون در الگوریتم‌های طبقه‌بندی و انتخاب ویژگی؛ ۴. شاخص دقت طبقه‌بندی برای ارزیابی الگوهای حاصل از الگوریتم طبقه‌بندی؛ ۵. شاخص دیویس بولدین<sup>۳۳</sup> برای بررسی میزان شباهت بین فرایندها در هر خوشه و تعیین تعداد مناسب خوشه‌ها؛ ۶. شاخص همبستگی ویژگی‌ها با ویژگی هدف در الگوریتم انتخاب ویژگی.

بعد از ارزیابی، پیشنهادی‌های بهبود فرایندها براساس الگوهای کشف‌شده ارائه می‌شوند. این الگوها باید مبتنی بر مفاهیم بهبود فرایندها باشند. یک تعامل بین داده‌کاو و کارشناس بهبود برای تحلیل الگوها و پیشنهادی‌های ارائه‌شده انجام می‌شود. الگوها باید به وسیله داده‌کاو و متخصص بهبود قابل به‌کارگیری باشد تا بتوان پیشنهادی‌های بهبود ارائه داد. این پیشنهادی‌ها باید با ادارات بهبود فرایندها و واقعیت‌های مرتبط به سازمان و محیط آن هم‌نوا باشد. بین خبرگان شامل متخصص بهبود، داده‌کاو، مدیران ارشد و کارشناسان فرآیند مشارکت انجام می‌شود. درنهایت پیشنهادی‌های بهبود با در نظر گرفتن همه موضوعات مرتبط به بافت سازمانی اجرا می‌شود. فرایندها با این پیشنهادی‌ها، تعدیل و اصلاح می‌شوند و در نتیجه ارتقا می‌یابند. فرایندهای جدید اجرا و عملکرد آنها ارزیابی می‌شود. این موضوع بررسی می‌شود که آیا تغییرات منتج از اجرای پیشنهادی‌ها با بافت سازمان و دنیای واقعی منطبق است یا خیر؟ بعد از اجرای فرایندهای بهبودیافته، مجموعه فرآیند و ویژگی‌ها برای اجرای مجدد چارچوب پیشنهادی به‌روزرسانی می‌شوند. با اجرای مجدد چارچوب پیشنهادی، الگوهای جدید شناسایی و با الگوهای قبلی جایگزین می‌شوند. این الگوها در برنامه بهبود مستمر و فزاینده به کار می‌روند. بدین ترتیب اجرای چارچوب پیشنهادی به صورت چرخه‌ای مستمر تکرار می‌شود.

## مطالعه موردی

### توضیح سازمان برای فهم کسب‌وکار

سازمان در حال مطالعه، پروژه محور و فناوری محور با تعداد ۱۲۰ نفر نیروی انسانی است. فعالیت اصلی سازمان برپایه پژوهش است. فرآیندها در واحدهای مدیریت بازرگانی، مدیریت پشتیبانی، مدیریت منابع انسانی، مدیریت مالی، مدیریت کیفیت، مدیریت دانش و آموزش، مدیریت عالی، واحدهای طراحی، واحد تست، کنترل پروژه، آینده پژوهی و مدیریت راهبردی، ساخت و بهره‌برداری، فناوری اطلاعات، مدیر پروژه و بازرسی اجرا می‌شوند.

برای شناسایی فرآیندها، وضعیت جاری سازمان مطالعه شده است. همچنین اهداف سازمان مرتبط با اهداف داده‌کاوی و اهداف بهبود فرآیند تعیین شدند. برای اجرای چارچوب پیشنهادی، تعاملی بین کارکنان اصلی سازمان، مدیران اجرایی، کارشناس بهبود و داده‌کاو انجام شده است. همچنین تحلیل کسب‌وکار برای ایجاد هم‌راستایی بین داده‌کاوی و بهبود فرآیند اجرا شده است. مسئله اصلی در سازمان بهبود فرآیند در حجم بسیار زیاد فرآیندها است. هدف، شناسایی الگو در حجم زیاد فرآیندها برای ارائه پیشنهادی بهبود فرآیند است. بدین منظور از چارچوب پیشنهادی برای به‌کارگیری داده‌کاوی در بهبود فرآیندها استفاده می‌شود. توجیحات به‌کارگیری هریک از الگوریتم‌های داده‌کاوی در بهبود فرآیندها در بخش‌های مرتبط مشاهده می‌شود.

### مجموعه فرآیندها

برای ارزیابی قابلیت اجرای چارچوب پیشنهادی از مجموعه فرآیندهای واقعی استفاده می‌شود. برای تهیه مجموعه فرآیندها، همکاری بین تحلیل‌گر فرآیند، داده‌کاو، کارشناس بهبود و مالک فرآیند انجام شده است. فرآیندها در یک بانک اطلاعاتی در قالب جدولی ذخیره شدند. بانک اطلاعاتی شامل آیتم‌هایی مانند ۱. تعداد زیاد نام فرآیند در ردیف‌های جدول؛ ۲. ویژگی‌های فرآیندها در ستون‌های جدول؛ ۳. مقادیر ویژگی‌های فرآیندها برای هر فرآیند در سلول مربوطه است. تعداد فرآیندها برابر با ۱۳۱۸ فرآیند است. فرآیندها با روش‌های زیر شناسایی و تهیه شده‌اند:

۱. ادبیات بهبود فرآیندها و موضوعات مرتبط؛ ۲. روش‌ها، دستورالعمل‌ها، مستندات، ماموریت، شرح وظایف واحدهای سازمانی، شرح مشاغل، شناسنامه مشاغل، شناسنامه فرآیندها؛ ۳. نظرات متخصص بهبود فرآیندها؛ ۴. روش طبقه‌بندی فرآیندها (برن، ۲۰۰۸)؛ ۵. مشاهده اجرای فرآیندها؛ ۶. چارچوب زنجیره ارزش پورتر؛ ۷. مجری فرآیند؛ ۸. مدل ارزیابی عملکرد واحدهای سازمان؛ ۹. مدل شایستگی منابع انسانی؛ ۱۰. کتابچه مهندسی مجدد ساختار سازمانی؛ ۱۱. اظهارنامه تعالی منابع انسانی؛ ۱۲. مراجع بالاسری؛ ۱۳. استانداردهای کیفیتی، فنی و تست پروژه‌ها؛ ۱۴. استانداردهای پیکره دانش مدیریت پروژه و اسناد فنی پروژه‌ها.

## مجموعه ویژگی‌های فرآیندها

بانک اطلاعاتی فرآیندها دارای ۸۰ ویژگی متنوع است. برای شناسایی ویژگی‌ها از همان مراجع شناسایی فرآیندها بهره گرفته شده است. جدول ۲ اطلاعات ویژگی‌های فرآیند را نشان می‌دهد؛ این اطلاعات شامل ۱. نام ویژگی، عنوانی که فرآیند را توصیف می‌کند؛ ۲. روش اندازه‌گیری مقادیر ویژگی‌ها که به‌طور نمونه شامل دو روش مشاهده اسناد و مدارک مرتبط به هر فرآیند یا دریافت نظرات است؛ ۳. شیوه دستیابی به مقدار ویژگی شامل: استخراج از ادبیات پژوهشی، مطالعه روش‌ها و مستندات مربوط به هر فرآیند، نظرات متخصص بهبود فرآیندها، نظرات مالک فرآیند، نظرات داده‌کاو، مشاهده اجرای فرآیندها است؛ ۴. مقدار ویژگی: تعیین‌کننده رفتار فرآیند در یک ویژگی است. این مقدار از طریق تعیین روش اندازه‌گیری و شیوه دستیابی به مقدار ویژگی به دست می‌آید. برخی عوامل دیگر که در تعیین مقدار ویژگی فرآیند دخیل هستند عبارتند از نخست، نیاز سازمان برای نحوه اندازه‌گیری یک ویژگی؛ دوم، زمان و منابع لازم (از جمله نیروی انسانی) برای اندازه‌گیری ویژگی فرآیند؛ سوم، سختی و راحتی تعیین مقدار ویژگی فرآیند به وسیله کارشناس مربوطه؛ ۵. فرد تعیین‌کننده مقدار ویژگی: فردی که در تعیین مقدار ویژگی برای یک فرآیند تصمیم می‌گیرد؛ ۶. مرجع شناسایی ویژگی: مرجعی که ویژگی فرآیند را تعریف می‌کند. در جدول ۲، اطلاعات چند ویژگی فرآیند به‌طور نمونه ارائه شده است. در جدول ۳ همه عناوین ویژگی‌های فرآیند ارائه شده است. این ویژگی‌ها دارای چهار نوع اسمی، ترتیبی، فاصله‌ای و نسبی هستند که نوع آنها در ستون "مقدار ویژگی" از جدول ۲ قرار داده شده است.

جدول ۲- اطلاعات درخصوص ویژگی‌های فرآیند

ردیف	نام ویژگی	روش اندازه‌گیری	شیوه دستیابی به مقدار ویژگی	مقادیر ویژگی	فرد تعیین‌کننده مقدار ویژگی	مرجع شناسایی ویژگی فرآیند
۱	نام فرآیند	مشاهده، مطالعه اسناد، اخذ نظرات	مشاهده، اخذ نظرات، مطالعه سند مربوطه	نام فرآیند	مالک فرآیند/ کارشناس فرآیندها	ادبیات پژوهشی
۲	نوع تأمین‌کننده	مشاهده، مطالعه اسناد، اخذ نظرات	مشاهده، مطالعه اسناد	واحد سازمانی	مالک فرآیند/ کارشناس فرآیندها	ادبیات پژوهشی
۳	نوع ورودی	مشاهده، مطالعه اسناد، اخذ نظرات	مشاهده، مطالعه اسناد	مستندات، منابع انسانی، اطلاعات، پول، ابلاغیه‌های بالاسری، درخواست، سیستم، قطعه‌کاری	مالک فرآیند/ کارشناس فرآیندها	ادبیات پژوهشی
۴	اهمیت فناوری	تعریف عملیاتی پژوهشگر: از نظر مالی، نیروی انسانی، سازمانی و فراسازمانی	نظر تخصصی مالک فرآیند و کارشناس توسعه فناوری	۱ الی ۵	کارشناس توسعه فناوری	ادبیات پژوهشی

### جدول ۳- عناوین ویژگی‌های فرآیند

ویژگی‌های فرآیند
<p>نام فرآیند؛ فرآیند نوشته‌شده یا نوشته‌نشده؛ نوع تأمین‌کننده؛ نوع ورودی؛ نوع خروجی؛ نوع مشتری؛ نوع مکانیزم؛ نوع کنترل؛ اهمیت فناوری؛ نوع فناوری؛ پیچیدگی؛ حیطة؛ مشروط بودن به عوامل تأثیرگذار محیطی؛ تعامل با بیرون سازمان؛ تأثیر بر سازمان‌های مرتبط دیگر؛ تأثیر بر کسب‌وکار؛ ارتباط مستقیم با پروژه‌های سازمان؛ تأثیر درخور توجه بر فرایندهای دیگر؛ تأثیرپذیری درخور توجه از فرایندهای دیگر؛ میزان پشتیبانی از دانش؛ میزان دارابودن قابلیت انتقال دانش بین افراد؛ میزان کپی‌برداری از فرآیند؛ قابلیت جایگزینی؛ میزان نادر بودن؛ نوع اتلاف؛ اهمیت استراتژیک؛ کلیدی بودن؛ برتری رقابتی؛ قابلیت برون‌سپاری؛ میزان درجه تکرارپذیری؛ تعداد بهبودهای انجام‌شده؛ ارزش مالی؛ هزینه؛ نوع فرآیند از نظر ارزش؛ محل اجرا؛ کارمند اصلی مرتبط؛ فرآیند مشتری‌مدار؛ فرآیند مباحثه‌ای؛ میزان رسمیت؛ درجه ساخت‌یافتگی؛ میزان خودکار بودن؛ سطح انتزاع؛ میزان نیاز به بازرسی و اندازه‌گیری؛ نوع تغییر؛ راحتی در اجرا/ امکان‌سنجی اجرا؛ میزان ریسک؛ ریسک شکست؛ جنس منابع؛ دسترسی منابع؛ سرعت اجرا؛ زمان لازم برای اجرا؛ نیاز به تخصص مدیریتی؛ میزان نیاز به مهارت برای اجرا؛ نیاز به قضاوت‌های انسانی (شخصی) و تجربیات؛ میزان نیاز به مطالعه کار؛ میزان نیاز به آموزش؛ میزان نیاز به نوآوری؛ میزان نیاز به توجه به کیفیت؛ ارتباط با فعالیت‌های مدیریت دانش؛ نوع کارشناس فرآیند؛ نیاز به فناوری اطلاعات؛ میزان قابلیت توانمندسازی کردن از طریق فناوری اطلاعات؛ میزان نیاز به جمع‌آوری اطلاعات؛ میزان نیاز به امنیت؛ نوع امنیت فرآیند؛ نیاز به مدیریت حالت؛ ارتباط مستقیم با اهداف و مأموریت سازمان؛ میزان نیاز به تبعیت از اهداف و مأموریت سازمان؛ هدف فرآیند؛ روش‌های بهبود فرآیند؛ امیک بودن یا نبودن؛ نوع فرآیند؛ رشته مرتبط؛ قابلیت پیش‌گویی؛ میزان سطح عدم‌اطمینان و ابهام.</p>

### ارزیابی اعتبار مجموعه فرآیندها

روش‌های ارزیابی اعتبار مجموعه فرآیندها عبارتند از: ۱. به‌کارگیری مجموعه فرآیندهای واقعی؛ ۲. تنوع وسیعی از مراجع برای تهیه بانک اطلاعاتی فرآیندها؛ ۳. ادبیاتی وسیع برای شناسایی ویژگی‌ها (جدول ۳)؛ ۴. تعریف عملیاتی ویژگی‌ها؛ ۵. اخذ نظرات خبرگان درخصوص مقادیر ویژگی‌ها.

### روش‌های آماده‌سازی مجموعه فرآیندها

در داده‌کاوی به‌کارگیری روش‌های آماده‌سازی و پیش‌پردازش داده‌ها باعث دستیابی به نتایج بهتر می‌شود. لارُس (۲۰۰۵) تعداد زیادی از این روش‌ها را توضیح داده است. برخی از روش‌های استفاده‌شده در این مطالعه عبارتند از:

۱. ادغام مجموعه فرآیندها از چند بانک اطلاعاتی برای ایجاد بانک اطلاعاتی یکپارچه؛ ۲. تغییر نوع ویژگی‌ها؛ برای مثال تبدیل نوع ویژگی از نسبی به ترتیبی؛ ۳. حذف فرآیندهای تکراری با مفاهیم یکسان؛ ۴. حذف ویژگی‌های تکراری با معانی یکسان؛ ۵. حذف ویژگی‌ها با مقادیر منحصر به فرد؛ ۶. حذف ویژگی‌ها با مقادیر یکسان؛ ۷. اصلاح اسامی ویژگی‌ها برای تخصیص مقادیر مناسب‌تر به آنها.

### یافته‌های پژوهش

این بخش یافته‌های به‌کارگیری چارچوب پیشنهادی پژوهش را تبیین می‌کند. برای ارزیابی اثربخشی و کاربردپذیری چارچوب پیشنهادی، چند نمونه از به‌کارگیری الگوریتم‌های داده‌کاوی برای استخراج الگوهای پنهان در فرآیندها ارائه می‌شود. این نمونه‌ها مبتنی بر ادبیات بهبود فرآیندها هستند. در نهایت با استفاده از الگوهای مستخرج، پیشنهادها به‌بود فرآیند ارائه می‌شود. الگوریتم‌های داده‌کاوی به‌وسیله نرم‌افزار کلمنتاین نسخه ۱۲ اجرا شده است.

## طبقه‌بندی

با استفاده از الگوریتم طبقه‌بندی، فرآیندها براساس یک ویژگی هدف طبقه‌بندی می‌شود. خروجی طبقه‌بندی به دو صورت است: ۱. قوانین اگر-آنگاه؛ ۲. ویژگی‌های مهمی که نسبت به سایر ویژگی‌ها بهتر می‌توانند فرآیندها را براساس ویژگی هدف طبقه‌بندی کنند.

در اینجا ذکر دو نکته اهمیت دارد؛ نخست اینکه الگوریتم درخت تصمیم‌گیری سی‌پنج از شاخص کسب اطلاعات برای انتخاب ویژگی‌های مؤثر بر ویژگی هدف استفاده می‌کند؛ یعنی هر ویژگی که بتواند قدرت کسب اطلاعات بیشتری داشته باشد و ویژگی هدف را بهتر تبیین کند، یک ویژگی مهم است و انتخاب می‌شود. از روش تحلیل همبستگی آماری برای انتخاب ویژگی‌ها در الگوریتم درخت تصمیم‌گیری استفاده می‌شود؛ اما روش‌های داده‌کاوی در مقایسه با آمار باتوجه به حجم بسیار زیاد داده، به‌طور هم‌زمان اثرات چند ویژگی بر ویژگی هدف را تعیین می‌کنند و این امر به‌صورت غیرخطی انجام می‌شود؛ یعنی در شناسایی چند متغیر به‌طور هم‌زمان روابط خطی حاکم نیست.

دوم اینکه با استفاده از الگوریتم طبقه‌بندی درخت تصمیم‌گیری ویژگی‌های مؤثر بر ویژگی هدف انتخاب می‌شود. به عبارت دیگر سه نوع روش انتخاب ویژگی وجود دارد؛ این روش‌ها عبارتند از: ۱. روش فیلتر<sup>۴۴</sup>؛ ۲. روش راپر<sup>۴۵</sup>؛ ۳. روش جاسازی‌شده<sup>۴۶</sup>. در روش فیلتر ارتباط بین ویژگی‌ها با ویژگی هدف بررسی می‌شود. الگوریتم انتخاب ویژگی مطرح‌شده از این روش استفاده می‌کند. در روش راپر ویژگی‌ها به‌گونه‌ای انتخاب می‌شوند که بتوانند رکوردها را براساس ویژگی هدف طبقه‌بندی کنند. در روش جاسازی‌شده اگر یک الگوریتم طبقه‌بندی اجرا شود، خروجی این الگوریتم دارای تعدادی ویژگی است. این ویژگی‌ها می‌توانند برای ویژگی‌های مؤثر بر ویژگی هدف انتخاب شوند. در این مقاله درخت تصمیم‌گیری ساخته‌شده دارای تعدادی ویژگی است. این ویژگی‌ها، ویژگی‌های مهمی هستند که شناسایی و انتخاب می‌شوند.

پیشنهاد‌های بهبود براساس دو خروجی طبقه‌بندی ارائه می‌شود؛ این پیشنهادها عبارتند از ۱. بعد از اجرای الگوریتم درخت تصمیم‌گیری، تعدادی قوانین اگر-آنگاه به‌وجود می‌آید. ویژگی‌هایی که این قوانین را می‌سازند ویژگی‌های مهمی هستند که برای ارائه پیشنهاد‌های بهبود انتخاب می‌شوند. همچنین در این قوانین، مقادیر ویژگی‌های فرآیندها براساس ویژگی هدف مشخص شده است. این مقادیر در ارائه پیشنهاد‌های بهبود فرآیندها کمک‌کننده هستند؛ ۲. تعداد ۱۰ ویژگی مهم انتخاب می‌شوند که بر ویژگی هدف تأثیر دارند. با استفاده از این ویژگی‌ها پیشنهاد‌های بهبود ارائه می‌شود.

تعداد ویژگی‌ها بستگی به نظر کارشناس بهبود فرآیندها و کارشناس داده‌کاوی دارد. در اینجا برای نمونه تعداد ۱۰ ویژگی انتخاب شده است. همچنین الگوریتم‌های درخت تصمیم‌گیری و انتخاب ویژگی، ویژگی‌ها را براساس اهمیت آن‌ها اولویت‌بندی می‌کنند. در تعیین تعداد ویژگی‌ها برخی نکات اهمیت دارد (نکته: مهم‌ترین ویژگی‌ها برای شناسایی و بهبود فرآیند انتخاب می‌شود؛ زیرا همه ویژگی‌های فرآیند مهم نیستند). ۱. هرچه تعداد ویژگی‌های فرآیند بیشتر باشد، تعداد بیشتری از ویژگی‌های مهم برای شناسایی فرآیندها انتخاب می‌شود. ۲. اگر تعداد ویژگی‌ها کم باشد، تعداد کمتری ویژگی مهم برای بهبود فرآیندها انتخاب می‌شود؛ ۳. برای تعیین تعداد ویژگی‌ها (باتوجه به اینکه ویژگی‌ها به‌لحاظ اهمیت رتبه‌بندی شده‌اند) براساس روش قضاوتی تعدادی از ویژگی‌های با رتبه بالا انتخاب می‌شوند؛ به‌گونه‌ای که این ویژگی‌ها براساس نظرات کارشناس بهبود فرآیند و داده‌کاوی، ارتباط بیشتری با ویژگی هدف دارند و بهتر می‌توانند ویژگی هدف را تبیین کنند.

در ادامه نمونه‌ای از به‌کارگیری درخت تصمیم‌گیری برای ارائه پیشنهاد‌های بهبود تشریح می‌شود.

در این مقاله، ویژگی هدف «کلیدی بودن فرآیند» برای نمونه انتخاب شده است و فرآیندها بر این اساس طبقه‌بندی می‌شوند. ویژگی‌های هدف متنوعی برای انتخاب ویژگی‌ها در چارچوب پیشنهادی در نظر گرفته می‌شود؛ اما انتخاب ویژگی هدف بستگی به نیاز سازمان در بهبود فرآیندها دارد. انتخاب ویژگی هدف با هماهنگی بین متخصص بهبود فرآیند و داده‌کاو انجام می‌شود.

برای طبقه‌بندی فرآیندها از درخت تصمیم‌گیری سی‌پنج استفاده شده است. دقت طبقه‌بندی مدل در مجموعه داده آموزش و آزمون به ترتیب برابر با ۹۰/۸۲ و ۹۲/۳۱ درصد است. جدول ۴ نتایج اعتبارسنجی متقاطع در داده‌های آموزش و آزمون را ارائه می‌دهد.

جدول ۴- اعتبارسنجی متقاطع

مجموعه داده آموزش	۱ (مثبت)	صفر (منفی)
۱ (مثبت)	مثبت درست: ۴۵۷	مثبت نادرست: ۸۲
صفر (منفی)	منفی نادرست: ۲۷	منفی درست: ۶۲۲
مجموعه داده آزمون	۱ (مثبت)	صفر (منفی)
۱ (مثبت)	مثبت درست: ۶۱	مثبت نادرست: ۹
صفر (منفی)	منفی نادرست: ۱	منفی درست: ۵۹

در جدول ۵ نمونه قوانین ایجادشده در طبقه‌بندی فرآیندهای کلیدی ارائه شده است. فرآیندهای کلیدی با این قوانین طبقه‌بندی می‌شوند. همچنین همه فرآیندهای سازمان براساس این قوانین طبقه‌بندی می‌شوند. اگر هدف بررسی فرآیند جدیدی باشد با قراردادن ویژگی‌های آن فرآیند در این قوانین می‌توان به کلیدی بودن یا نبودن آن فرآیند پی برد. همچنین برای شناسایی فرآیندهای کلیدی به الگوها و قوانینی دسترسی وجود دارد که با استفاده از آنها کلیدی بودن یک فرآیند تشخیص داده می‌شود. جدول ۵ بیان می‌کند که برای شناسایی فرآیندهای کلیدی کافی است چند ویژگی آن بررسی شود و لازم نیست همه ویژگی‌ها بررسی شود؛ بنابراین با سرعت بیشتر و راحت‌تر، فرآیندهای کلیدی شناسایی می‌شود.

کاربرد این قوانین در چارچوب پیشنهادی این است که نخست می‌توان با قوانین، فرآیندهای کلیدی را شناسایی کرد و برای شناسایی آنها نیازی نیست همه ویژگی‌ها بررسی شوند. دوم اینکه با استفاده از این قوانین، مشخصات فرآیندها شناسایی می‌شود. درواقع شناسایی مشخصات فرآیندها در ارائه پیشنهادها بهبود فرآیند کمک‌کننده است. به عبارت دیگر پیشنهادات بهبود فرآیند هدفمند و از ارائه پیشنهادات بهبود به صورت غیرهدفمند جلوگیری می‌شود. از طرف دیگر با تغییر ویژگی هدف، فرآیندها از نقطه‌نظر ابعاد دیگر شناسایی می‌شوند و پیشنهادات بهبود فرآیند به طور متناسبی ارائه می‌شود.

مهم‌ترین ویژگی‌های فرآیندی در ساخت درخت تصمیم‌گیری به ترتیب اولویت و اهمیت در جدول ۶ مشخص است. این ویژگی‌ها به صورت نمونه هستند و می‌توان از سایر ویژگی‌های مؤثر بر ویژگی هدف برای ارائه پیشنهادها بهبود فرآیند استفاده کرد. بعد از استخراج قوانین اگر-آنگاه و مهم‌ترین ویژگی‌ها در ساخت درخت تصمیم‌گیری، پیشنهادهایی برای بهبود فرآیندهای کلیدی ارائه می‌شود. همان طور که در جدول ۶ مشاهده می‌شود، پیشنهاد بهبود به طور خاص و فقط براساس ویژگی فرآیندی مدنظر ارائه شده است. ویژگی‌های فرآیندی

در جدول ۶ همان ویژگی‌هایی هستند که در ساخت چهار نمونه قانون (جدول ۵) به کار رفته‌اند؛ بنابراین پیشنهادهای بهبود با دو خروجی طبقه‌بندی، یعنی براساس نمونه قوانین اگر- آنگاه و براساس نمونه ویژگی‌های مؤثر بر ویژگی هدف ارائه شده است.

جدول ۵- نمونه قوانین در مدل طبقه‌بندی فرآیندهای کلیدی

قانون
اگر تأثیر فرآیند بر فرآیندهای دیگر درخور توجه باشد و میزان نادر بودن فرآیند از متوسط تا خیلی زیاد و ارزش مالی فرآیند زیاد و خیلی زیاد باشد، فرآیند کلیدی است. نمونه این فرآیند: جذب منابع مالی برای پروژه
اگر تأثیر فرآیند بر فرآیندهای دیگر درخور توجه باشد و میزان نادر بودن فرآیند از متوسط تا خیلی زیاد، نیاز به نوآوری در فرآیند از متوسط تا خیلی زیاد و نوع فرآیند فنی باشد، فرآیند کلیدی است. نمونه این فرآیند: تعیین سطح بلوغ فناوری‌ها
اگر نوع تأمین‌کننده فرآیند، دانش و ارزش مالی فرآیند از متوسط تا خیلی زیاد باشد، فرآیند کلیدی است. نمونه این فرآیند: تدوین سند راهبردی مدیریت دانش
اگر تأثیر فرآیند بر فرآیندهای دیگر درخور توجه باشد یا میزان نادر بودن فرآیند از متوسط تا خیلی زیاد و میزان قابلیت توانمندسازی کردن از طریق فناوری اطلاعات از خیلی کم تا متوسط باشد، فرآیند کلیدی است. نمونه این فرآیند: تأیید نمونه آزمایشی

جدول ۶- پیشنهادهای بهبود فرآیندهای کلیدی

ویژگی فرآیندی	پیشنهاد بهبود فرآیند
میزان نادر بودن	برنامه‌های توسعه نوآوری برای افزایش میزان نادر بودن فرآیند نسبت به رقبا
تأثیر درخور توجه بر فرآیندهای دیگر	۱. بهبود فرآیندهایی که از فرآیندهای کلیدی تأثیرپذیر هستند؛ ۲. شناسایی عواملی از فرآیندها که فرآیندهای کلیدی بر آنها تأثیر می‌گذارند.
ارزش مالی	۱. شناسایی و مدیریت عواملی که باعث افزایش ارزش مالی فرآیندهای کلیدی می‌شود؛ ۲. تهیه اقدامات و برنامه‌ها برای افزایش ارزش مالی فرآیندهای کلیدی؛ ۳. حذف، ادغام، یا برون‌سپاری فرآیندهایی که ارزش مالی ایجاد نمی‌کنند.
نوع فرآیند از نظر ارزش	شناسایی و بهبود فرآیندهای ارزش‌افزا و غیرارزش‌افزا در سازمان
تأثیر بر کسب‌وکار	۱. شناسایی و بهبود فرآیندهای تأثیرگذار بر کسب‌وکار؛ ۲. شناسایی و مدیریت شاخص‌های تأثیرگذار فرآیند بر کسب‌وکار.
نوع تأمین‌کننده	شناسایی و توجه به تأمین‌کنندگان (از جمله منابع انسانی و دانش) فرآیندهای کلیدی
تأثیرپذیری درخور توجه از فرآیندهای دیگر	۱. شناسایی و بهبود فرآیندهایی که بر فرآیندهای کلیدی تأثیر می‌گذارند؛ ۲. شناسایی و مدیریت عوامل و شاخص‌های تأثیرگذار بر فرآیندهای کلیدی.
دسترسی به منابع	۱. شناسایی روش‌های دسترسی آسان به منابع برای اجرای بهتر فرآیندهای کلیدی؛ ۲. شناسایی و مدیریت منابع فرآیندهای کلیدی.
میزان نیاز به مهارت برای اجرا	توسعه مهارت و آموزش برای اجرای فرآیندهای کلیدی
کارمند اصلی مرتبط	۱. آموزش کارکنان و توسعه فردی آنها برای اجرای فرآیندهای کلیدی؛ ۲. ارزیابی شایستگی شغلی، جانشین‌پروری، حفظ و نگهداشت کارکنان فرآیندهای کلیدی.
میزان نیاز به نوآوری	۱. تدوین راهبرد مدیریت نوآوری برای فرآیندهای کلیدی؛ ۲. به‌کارگیری کارکنان نوآور برای انجام فرآیندهای کلیدی.
میزان قابلیت توانمندسازی کردن با فناوری اطلاعات	۱. کاهش میزان رسمیت فرآیندهای کلیدی؛ ۲. به‌کارگیری فناوری اطلاعات برای اجرای فرآیندهای غیرکلیدی.



## خوشه‌بندی

با استفاده از الگوریتم خوشه‌بندی کامیانگین، فرایندها بخش‌بندی و پیشنهادهای بهبود متناسب با هر بخش تدوین می‌شود. برای تعیین تعداد مناسب خوشه‌ها از شاخص دیویس بولدین استفاده شده است. هر قدر مقدار این شاخص کمتر باشد، خوشه‌بندی با کیفیت بهتری انجام می‌شود. مقدار این شاخص در تعداد ۱۰ خوشه به کمترین مقدار برابر با  $1/65$  - رسیده است. همچنین در الگوریتم خوشه‌بندی، تعداد تکرارها تا رسیدن به کمترین خطا برابر با ۲۰ تکرار و کمترین خطا برابر با  $0/057$  است.

خروجی خوشه‌بندی، پروفایل خوشه‌ها است که فرایندها را در هر خوشه براساس ویژگی‌های آنها توصیف می‌کند. پروفایل خوشه‌ها در جدول ۷ نمایش داده شده است.

جدول ۷- پروفایل خوشه‌ها برای توصیف فرایندها

شماره خوشه	تعداد فرایندها در خوشه (تعداد)	میزان درجه تکرارپذیری (تعداد)	فرآیند مشتری‌مدار (درصد)	ارتباط مستقیم با پروژه‌های سازمان (درصد)	کلیدی بودن (درصد)	میزان دارابودن قابلیت انتقال دانش بین افراد (درصد)	درجه ساخت‌یافتگی (درصد)
۱	۲۰۰	۵۰	۱/۵۰	۱۱/۵۰	۸/۵۰	۱۸	۹۵
۲	۱۲۴	۶	۸۷/۷۱	۹۹/۱۹	۱۰۰	۹۹/۱۹	۳/۲۳
۳	۱۲۵	۷	۴۴	۶۳/۲۰	۹۶	۶۰	۴۰/۸۰
۴	۸۹	۵	۴/۴۹	۴۴/۹۴	۸۷/۶۴	۹۳/۲۶	۱۷/۹۸
۵	۱۸۸	۱۱	۲۳/۴۰	۴۱/۶۲	۷۰/۲۱	۳۷/۲۳	۷۴/۴۷
۶	۱۹۱	۲۲۲	۰	۱/۵۷	۰/۵۲	۲/۰۹	۱۰۰
۷	۲۷	۳	۲۲/۲۲	۲۹/۶۳	۳۷/۰۴	۱۱/۱۱	۹۶/۳۰
۸	۱۵۸	۱۰	۴۳/۰۴	۹۴/۹۴	۸۴/۱۸	۹۴/۹۴	۵۳/۱۶
۹	۵۷	۲	۲۴/۵۶	۲۴/۵۶	۹۸/۲۵	۷۳/۶۸	۵/۲۶
۱۰	۱۵۹	۳۸	۸/۸۱	۳۹/۶۲	۲۳/۹۰	۱۶/۹۸	۸۹/۹۴

در ادامه فرایندهای خوشه نخست به‌طور نمونه توصیف می‌شود. در خوشه نخست، تعداد ۲۰۰ فرآیند وجود دارد. میانگین تکرارپذیری هر فرآیند برابر با ۵۰ تکرار در سال است.  $1/5$  درصد از فرایندهای این خوشه، مشتری‌محور (منظور مشتری بیرونی) هستند.  $11/50$  درصد از فرایندها ارتباط مستقیم با پروژه‌های سازمان دارند.  $8/5$  درصد فرایندهای این خوشه کلیدی هستند.  $18$  درصد از فرایندهای این خوشه دارای قابلیت انتقال دانش بین کارکنان هستند.  $95$  درصد از فرایندها درجه ساخت‌یافتگی زیادی دارند. چند نمونه از فرایندهای خوشه نخست عبارتند از محاسبات راندمان کارکنان در پروژه‌ها؛ انجام محاسبات حقوق کارکنان؛ ثبت اعتبارات و منابع پروژه‌ها؛ تهیه بانک اطلاعاتی شرکت‌های فناور؛ تعیین سختی کار مشاغل.

روش ارائه پیشنهاد بهبود به این صورت است که با استفاده از الگوریتم کامیانگین، فرایندها به چند خوشه تقسیم می‌شوند. پیشنهادهای بهبود بر دو اساس به‌طور هدفمند برای هر خوشه ارائه می‌شود که شامل ۱. پروفایل خوشه‌ها است که مشخصات فرایندها را در هر خوشه توصیف می‌کند؛ ۲. ویژگی‌های منتخب برای توصیف رفتار

فرآیندها در هر خوشه است. ویژگی‌های منتخب ویژگی‌هایی است که کارشناس بهبود فرآیندها و داده‌کاو برای خوشه‌بندی فرآیندها انتخاب کرده‌اند. در جدول ۸ پیشنهاد‌های بهبود برای فرآیندهای خوشه نخست (به‌طور نمونه) ارائه شده است.

جدول ۸- پیشنهاد‌های بهبود فرآیندهای خوشه نخست

ویژگی فرآیندی	پیشنهاد بهبود فرآیند
میزان درجه تکرارپذیری	خودکارسازی اجرای فرآیندها به‌کمک فناوری اطلاعات به‌دلیل تکرار زیاد فرآیند
فرآیند مشتری‌مدار	شناسایی فرآیندهای مشتری‌مدار که متأثر از فرآیندهای خوشه نخست هستند و طراحی سیستم ارزیابی عملکرد فرآیندهای خوشه نخست برای بررسی کیفیت اجرای این فرآیندها به‌عنوان ورودی فرآیندهای مشتری‌مدار
ارتباط مستقیم با پروژه‌های سازمان	ادغام فرآیندهای این خوشه با فرآیندهای مرتبط به پروژه برای افزایش غنای فرآیندها
کلیدی بودن	۱. حذف فرآیندهای غیرکلیدی. ۲. حذف فعالیت‌های غیرکلیدی فرآیندها در خوشه نخست.
میزان دارابودن قابلیت انتقال دانش بین افراد	ترکیب فرآیندهای این خوشه با فرآیندهایی که قابلیت زیادی در انتقال دانش دارند
درجه ساخت‌یافتگی	تدوین روش و دستورالعمل برای فهم اجرای فرآیند به‌وسیله کارکنان

وقتی الگوریتم خوشه‌بندی اعمال می‌شود، این الگوریتم نمی‌تواند همه فرآیندها را با در نظر گرفتن دقیق همه ویژگی‌های فرآیند بخش‌بندی کند. مسلماً در هر خوشه، فرآیندهایی قرار می‌گیرند که ممکن است در یک ویژگی با فرآیند دیگر در همان خوشه مشابه نباشند. در خوشه‌بندی سعی می‌شود فرآیندهای موجود در یک خوشه حداکثر شباهت را به هم داشته باشند؛ بنابراین ممکن است در یک خوشه فرآیندهایی باشند که در یک یا چند ویژگی مشابه با فرآیندهای دیگر عمل نمی‌کنند؛ برای مثال در خوشه نخست تعداد سه فرآیند مشتری‌مدار وجود دارد. لزوماً پیشنهادات بهبود فرآیند برای همه فرآیندها در یک خوشه اعمال نمی‌شود؛ بلکه بررسی مبتنی‌بر روش قضاوتی، با توجه به محیط سازمان و ادراکات بهبود فرآیند و براساس نظرات مالک فرآیند، کارشناس بهبود و داده‌کاو انجام می‌شود. همچنین قابلیت اجرایی بودن پیشنهاد بهبود ارزیابی می‌شود. پیشنهادات بهبود مطرح در جدول ۸ نمونه‌ای از پیشنهاداتی هستند که در مقاله آورده شده است. برای اجرای این پیشنهادات باید به تمام نکات مطرح شده توجه شود؛ اما در ارائه پیشنهاد بهبود فرآیند برای فرآیندهای یک خوشه، سعی می‌شود پیشنهاد به‌گونه‌ای مطرح شود که برای بیشتر فرآیندهای آن خوشه اجراشدنی باشد. مسلماً برای برخی فرآیندها که مشابه فرآیندهای دیگر در همان خوشه هستند، این پیشنهادها اعمال نمی‌شود و پیشنهاد‌های دیگری در نظر گرفته می‌شود.

### انتخاب ویژگی‌ها

با استفاده از الگوریتم انتخاب ویژگی، ویژگی‌های مهم فرآیندها شناسایی و سپس پیشنهاد‌های بهبود براساس آنها ارائه می‌شود. روش ارائه پیشنهاد بهبود به این شکل است که ابتدا یک ویژگی هدف با تعامل بین کارشناس بهبود و داده‌کاو تعیین می‌شود. در این مقاله ویژگی هدف "هزینه فرآیند" برای نمونه انتخاب شده است. ویژگی‌های هدف متنوعی برای انتخاب ویژگی‌ها در چارچوب پیشنهادی در نظر گرفته می‌شود. انتخاب ویژگی هدف با توجه به نیاز سازمان در بهبود فرآیند و براساس تعامل بین متخصص بهبود فرآیند و داده‌کاو انجام می‌شود.

الگوریتم انتخاب ویژگی، ویژگی‌های فرآیندی مهم را براساس ویژگی هدف انتخاب می‌کند. در مجموعه ویژگی‌ها، ویژگی‌های منتخب همبستگی بیشتری با ویژگی هدف دارند. کمترین ضریب تغییر<sup>۴۷</sup> و انحراف استاندارد<sup>۴۸</sup> در الگوریتم انتخاب ویژگی به ترتیب برابر با ۰/۱ و صفر است. با استفاده از ویژگی‌های منتخب، ارائه هدفمندی از پیشنهادها به‌طور انجام می‌شود؛ به عبارت دیگر پیشنهادها به‌طور فقط براساس مهم‌ترین ویژگی‌های منتخب مستخرج از اجرای الگوریتم انتخاب ویژگی ارائه می‌شود؛ برای نمونه جدول ۹ تعداد ۱۰ ویژگی مهم را ارائه می‌کند که با ویژگی هدف "هزینه فرآیند" همبستگی بیشتری دارند. این ۱۰ ویژگی نسبت به سایر ویژگی‌ها توانایی بیشتری برای تبیین هزینه فرآیند دارند. اگر سازمان بخواهد پیشنهاد بهبود فرایندها را برای کاهش هزینه فرآیند ارائه دهد، بهتر است به این ۱۰ ویژگی توجه کند. جدول ۹، پیشنهادها به‌طور براساس ۱۰ ویژگی مهم منتخب تأثیرگذار بر هزینه فرآیند ارائه می‌دهد.

جدول ۹- پیشنهادها به‌طور برای کاهش هزینه فرآیند

ویژگی فرآیندی	پیشنهاد بهبود فرآیند
اهمیت فناوری	۱. مدیریت فناوری فرآیند و انتخاب یک فناوری که هزینه زیادی به سازمان تحمیل نمی‌کند. ۲. تعامل با شرکت‌های فناوری براساس برنامه توسعه فناوری.
نوع مکانیزم	تحلیل مکانیزم اجرای فرآیند (از جمله پول، نیروی انسانی، و فناوری)
ریسک شکست	مدیریت ریسک فرایندها
جنس منابع	مدیریت منابع فرایندها
برتری رقابتی	شناسایی و توسعه فرآیندهایی که باعث افزایش برتری رقابتی می‌شوند
میزان کپی‌برداری از فرآیند	۱. حذف، برون‌سپاری، یا ادغام فرآیندهایی که کپی‌برداری از آنها راحت است؛ ۲. آموزش کارکنان با هدف توسعه فرآیندهایی که کپی‌برداری از آنها سخت است
نوع فرآیند	شناسایی نوع فرآیند با هدف مدیریت آن در واحد سازمانی برای کاهش هزینه فرآیند
نوع فرآیند از نظر ارزش	۱. تحلیل هزینه-ارزش برای هر فرآیند به‌خصوص فرآیندهای هزینه‌زا؛ ۲. شناسایی و بهبود فرآیندهایی که ارزش کم ایجاد می‌کنند، اما هزینه زیادی دارند.
نوع اتلاف	شناسایی و تعیین نوع اتلاف برای فرآیندهای هزینه‌زا
ارزش مالی	تحلیل هزینه و ارزش مالی برای فرآیندهای هزینه‌زا

### بحث و نتیجه‌گیری

این مقاله چارچوبی برای به‌کارگیری روش‌های داده‌کاوی در حجم زیاد فرایندها با هدف شناسایی الگوهای مفید برای بهبود فرایندها ارائه می‌دهد. مجموعه واقعی از فرایندها همراه با ویژگی‌های آنها برای ارزیابی کاربردپذیری و اثربخشی چارچوب پیشنهادی به کار گرفته شد. ویژگی‌های فرایندها براساس مطالعه در مفاهیم بهبود فرایندها و مدیریت دانش شناسایی شدند.

در چارچوب پیشنهادی بعد از تهیه مجموعه فرایندها و ویژگی‌های آنها با استفاده از روش‌های داده‌کاوی، الگوهای ارزشمند شناسایی شد. از این الگوها برای ارائه پیشنهادها به‌طور استفاده شد. روش‌های داده‌کاوی به شرح زیر است:

نخست به‌کارگیری درخت تصمیم‌گیری برای طبقه‌بندی فرایندها براساس ویژگی هدف. دوم به‌کارگیری

الگوریتم کامیاب‌ترین برای بخش‌بندی فرآیندها. سوم به‌کارگیری روش انتخاب ویژگی برای شناسایی ویژگی‌های مهم. روش‌های بهبود فرآیندها، فقط فرآیندها با تعداد ویژگی‌های کم را بهبود می‌دهند. وقتی تعداد فرآیندها و ویژگی‌های آنها بسیار زیاد شود، این روش‌ها نمی‌توانند به‌طور سریع و با دقت زیاد پیشنهادها را بهبود ارائه دهند. در چارچوب پیشنهادی، به‌دلیل وجود تعداد زیاد ویژگی‌های فرآیندها، پیشنهادها را بهبود می‌دهند و تنوع وسیع‌تری هستند؛ اما روش‌های بهبود فرآیندها، پیشنهادها را براساس تعداد محدودی ویژگی ارائه می‌دهند. از طرفی مطالعات اندکی روی کاربرد داده‌کاوی در بهبود فرآیندها وجود دارد. این مطالعات دارای برخی تفاوت‌ها نسبت به چارچوب پیشنهادی هستند که در جدول ۱۰ ارائه شده است.

جدول ۱۰ مقایسه چارچوب پیشنهادی با مطالعات قبلی به‌کارگیری داده‌کاوی در بهبود فرآیندها

مطالعات قبلی کاربرد داده‌کاوی در بهبود فرآیندها	چارچوب پیشنهادی
نبود فهم کافی از حجم زیاد فرآیندها و تصویری جامع و یکپارچه از اطلاعات آنها	به‌کارگیری مجموعه وسیعی از فرآیندهای واقعی برای استخراج الگوهای پنهان ارزشمند برای بهبود فرآیندها
داشتن نگرشی مجزا و ایزوله‌شده به فرآیندها	به‌کارگیری نگرشی یکپارچه به همه فرآیندها در سازمان
به‌کارگیری تعداد محدودی از ویژگی‌های فرآیندی	به‌کارگیری تنوع وسیعی از ویژگی‌های فرآیندی
به‌کارگیری یک روش داده‌کاوی برای تحلیل فرآیندها	به‌کارگیری تنوعی از روش‌های داده‌کاوی در تحلیل فرآیندها
توصیف روابط تئوریک، عمومی و کلی بین داده‌کاوی و بهبود فرآیندها و توجه کم به یکپارچگی بین آنها	لحاظ‌کردن نتایج مستخرج از به‌کارگیری داده‌کاوی برای بهبود فرآیندها در قالب چارچوب پیشنهادی
ارائه پیشنهادها با تنوع کم	ارائه تنوع وسیعی از پیشنهادها با تنوع فرآیند

برخی اقدامات برای ارزیابی اعتبار چارچوب پیشنهادی انجام شد که عبارتند از ۱. به‌کارگیری مجموعه‌ای واقعی از فرآیندها؛ ۲. روش‌های اعتبارسنجی متقاطع و دیویس بولدین برای ارزیابی خروجی‌های داده‌کاوی؛ ۳. ارائه چند نمونه از به‌کارگیری الگوریتم‌های داده‌کاوی در استخراج الگوهای پنهان در فرآیندها برای ارزیابی اثربخشی و کاربردپذیری چارچوب پیشنهادی؛ ۴. ارزیابی اعتبار پیشنهادها به‌وسیله نظرات خبرگان و مفاهیم بهبود فرآیندها.

چارچوب پیشنهادی دارای برخی محدودیت‌ها در اجرا است. نخست اینکه جمع‌آوری و پیش‌پردازش مجموعه بزرگی از فرآیندها به‌همراه ویژگی‌های آنها سخت و زمان‌بر است. دوم اینکه تعیین مقادیر برای برخی ویژگی‌ها نیازمند دریافت نظرات خبرگان است. این موضوع نیاز به آموزش خبرگان و ایجاد زبان مشترک بین اهداف داده‌کاوی و اهداف بهبود فرآیند دارد. سوم اینکه پیشنهادها به‌طور اساس خروجی به‌کارگیری داده‌کاوی در مجموعه فرآیندهای متعلق به سازمان در حال مطالعه است. چهارم، اجرای چارچوب پیشنهادی دارای برخی الزامات همانند روش‌شناسی‌های بهبود فرآیندها است. از جمله تعهد مدیریت ارشد؛ بودجه و زمان؛ نیروی انسانی متخصص؛ آموزش مدیران؛ زیرساخت؛ توانمندی مجری چارچوب در برقراری ارتباطات با کارکنان. پنجم، چارچوب پیشنهادی از داده‌های ثبت وقایع بهره نبرده است که در فرآیندکاوی به کار می‌رود. سازمان‌ها می‌توانند از چارچوب پیشنهادی برای بهبود فرآیندهای خود استفاده کنند. همچنین این چارچوب می‌تواند به روش‌شناسی‌های بهبود فرآیند و فرآیندکاوی در حجم زیاد فرآیندها کمک کند.

پیشنهادهای پژوهش‌های آتی شامل موارد زیر است:

نخست، به‌کارگیری سایر روش‌های داده‌کاوی در چارچوب پیشنهادی که باعث الگوهای جدیدی برای بهبود فرایندها می‌شود. دوم در مطالعات بعدی از نمونه‌های دیگر برگرفته از ادبیات بهبود فرایندها برای ارزیابی چارچوب پیشنهادی بهره گرفته شود. سوم سایر روش‌شناسی‌های بهبود فرآیند همچون شش‌سیگما و مهندسی مجدد فرایندها در چارچوب پیشنهادی به کار روند. چهارم چارچوب پیشنهادی به بهبود فرایندهای با وابستگی شدید به دانش بپردازد. پنجم چارچوب پیشنهادی با روش‌شناسی‌های مدیریت دانش یکپارچه شود. ششم چارچوب پیشنهادی برای یک سیستم پشتیبان تصمیم‌گیری با هدف بهبود فرایندها توسعه یابد.

## References

- Borrego, D., Barba, I. (2014). Conformance checking and diagnosis for declarative business process models in data-aware scenarios. *Expert Systems with Applications*, 41: 5340–5352.
- Brown, S. P. (2008). Business Processes and Business Functions: a new way of looking at employment. *Monthly Labor Review*, 131: 51-70.
- Castellanos, M., Casati, F., Dayal, U., & Shan, M. C. (2004). A comprehensive and automated approach to intelligent business processes execution analysis. *Distributed and Parallel Databases*, 16(3): 239-273.
- Chen, F. Z., & Wang, X. Z. (1999). An integrated data mining system and its application to process operational data analysis. *Computers & Chemical Engineering*, 23: S787-S790.
- Claes, J., Poels, G. (2014). Merging event logs for process mining: A rule based merging method and rule suggestion algorithm. *Expert Systems with Applications*, 41: 7291–7306.
- Ćurko, K., Varga, M., & Lončar, A. (2008). The support of business intelligence technology in Process and Business Engineering. *International Journal of Computers*, 2 (1): 8-14.
- Dalmaris, P., Tsui, E., Hall, B., & Smith, B. (2007). A framework for the improvement of knowledge-intensive business processes. *Business Process Management Journal*, 13 (2): 279-305.
- Damij, N., Damij, T. (2014). Business Process Approaches, in: Process Management: A Multi-Disciplinary Guide to Theory, Modeling, and Methodology. Springer Berlin Heidelberg, 45–60.
- Delgado, A., Weber, B., Ruiz, F., Garcia-Rodríguez de Guzmán, I., Piattini, M. (2014). An integrated approach based on execution measures for the continuous improvement of business processes realized by services. *Information and Software Technology*, 56: 134–162.
- D'heygere, T., Goethals, P.L.M., De Pauw, N. (2003). Use of genetic algorithms to select input variables in decision tree models for the prediction of benthic macroinvertebrates. *Ecological Modelling*, 160: 291–300.
- Dumas, M., La Rosa, M., Mendling, J., & Reijers, H. A. (2013). Fundamentals of business process management (Vol. 1, p. 2). Heidelberg: Springer.
- Folorunso, O., & Ogunde, A. O. (2005). Data mining as a technique for knowledge management in business process redesign. *Information management & computer security*, 13 (4): 274-280.
- Ghanadbashi, S., Khanabaei, M., & Abadeh, M. S. (2013). Applying data mining techniques to business process reengineering based on simultaneous use of two novel proposed approaches. *International Journal of Business Process Integration and Management*, 6 (3): 247-267.
- Ghattas, J., Soffer, P., Peleg, M. (2014). Improving business process decision making based on past experience. *Decision Support Systems*, 59: 93–107.
- Gómez-Pérez, J.M., Erdmann, M., Greaves, M., Corcho, O., Benjamins, R. (2010). A framework and computer system for knowledge-level acquisition, representation, and reasoning with process

- knowledge. *International Journal of Human-Computer Studies*, 68: 641–668.
- Grigori, D., Casati, F., Castellanos, M., Dayal, U., Sayal, M., & Shan, M. C. (2004). Business process intelligence. *Computers in industry*, 53 (3): 321-343.
- Gröger, C., Schwarz, H., & Mitschang, B. (2014). Prescriptive analytics for recommendation-based business process optimization. In Business Information Systems. *Springer International Publishing*: 25-37.
- Hammer, M. (2002). Process management and the future of six sigma. *Sloan Management Review*, 43 (2): 26-32.
- Harrington, H.J. (1991). The return of high performance to the US workplace. *Journal of Business Strategy*, 12 (4): 23-27.
- Harmon, P. (2007). Business process change. Morgan Kaufmann.
- Houy, C., Fettke, P., Loos, P., van der Aalst, W.M.P., Krogstie, J. (2011). Business Process Management in the Large. *Business and Information Systems Engineering*, 3: 385–388.
- Huang, Z., Lu, X., Duan, H. (2012). Resource behavior measure and application in business process management. *Expert Systems with Applications*, 39: 6458–6468.
- Jeong, H., Song, S., Shin, S., Rae Cho, B. (2008). Integrating data mining to a process design using the robust bayesian approach. *International Journal of Reliability, Quality and Safety Engineering*, 15: 441–464.
- Jung, J., Choi, I., & Song, M. (2007). An integration architecture for knowledge management systems and business process management systems. *Computers in industry*, 58 (1): 21-34.
- Koh, H.C., Low, C.K. (2004). Going concern prediction using data mining techniques. *Managerial Auditing Journal*, 19: 462–476.
- Larose, D.T. (2005). *Discovering Knowledge in Data, an Introduction to Data Mining*, 1st ed. John Wiley & Sons, New Jersey.
- Lee, S.J., Siau, K. (2001). A review of data mining techniques. *Industrial Management & Data Systems*, 101: 41–46.
- Lepmets, M., McBride, T., Ras, E. (2012). Goal alignment in process improvement. *Juornal of Systems and Software*, 85: 1440–1452.
- Mathew, S. and George, S. T. (2012). Implementation of data mining technique for BPR. *International Conference on Electrical Engineering and Computer Science. Trivendum*: 262-266.
- Nissen, M. E. (1999). Knowledge-based knowledge management in the reengineering domain. *Decision Support Systems*, 27 (1): 47-65.
- Pivk, A., Vasilecas, O., & Kalibatiene, D. (2014). Ontology and SOA Based Data Mining to Business Process Optimization. In Information System Development. *Springer International Publishing*: 255-268.
- Ranjbar Fard, M., AGhdasi, M., Albadvi, A., & Hassanzadeh, M. (2013). Knowledge Management Barriers Identification for the Four Kinds of Business Processes. *Journal of Information Technology Management*, 5 (1): 61-88. (in persian)
- Rupnik, R., Jaklic, J. (2009). The deployment of data mining into operational business processes, in: *Data Mining and Knowledge Discovery in Real Life Applications*. I-Tech Education and Publishing, Vienna, 373–388.
- Salappa, A., Doumpos, M. and Zopounidis, C. (2007). Feature selection algorithms in classification problems: an experimental evaluation. *Optimization Methods and Software*, 22: 199-212.
- Sohail, A., Dhanapal Durai Dominic, P. (2012). A gap between Business Process Intelligence and redesign process, in: 2012 International Conference on Computer & Information Science (ICCIS). IEEE, 136–142.
- Song, M., & Van der Aalst, W. M. (2008). Towards comprehensive support for organizational mining. *Decision Support Systems*, 46(1): 300-317.

- Tan, P. N., Steinbach, M., & Kumar, V. (2006). *Introduction to Data mining* (Vol. 1). Boston: Pearson Addison Wesley.
- Tiwari, A., Turner, C. J., & Majeed, B. (2008). A review of business process mining: state-of-the-art and future trends. *Business Process Management Journal*, 14 (1): 5-22.
- Vukšić, V.B., Bach, M.P., Popovič, A. (2013). Supporting performance management with business process management and business intelligence: A case analysis of integration and orchestration. *International Journal of Information Management*, 33: 613–619.
- Wegener, D., & Rüping, S. (2010). On integrating data mining into business processes. 13<sup>th</sup> international conference on Business Information Systems, *Springer Berlin Heidelberg*: 183-194.
- Yang, H., Park, M., Cho, M., Song, M., & Kim, S. (2014, October). A system architecture for manufacturing process analysis based on big data and process mining techniques. In *Big Data (Big Data)*, 2014 IEEE International Conference on (pp. 1024-1029). IEEE.
- Zhonghua, D. and Limei, S. (2008). Business processes reengineering based on data mining. proceeding of 2008 international conference on logistics engineering and supply chain, china: 164-169.

- 
- <sup>1</sup> Harrington
  - <sup>2</sup> Dalmaris
  - <sup>3</sup> Davenport, Hammer, and Champy
  - <sup>4</sup> Harmon
  - <sup>5</sup> Brown
  - <sup>6</sup> Ranjbar Fard
  - <sup>7</sup> Jung
  - <sup>8</sup> Chen, and Wang
  - <sup>9</sup> Jeong
  - <sup>10</sup> Houy
  - <sup>11</sup> Lepmets
  - <sup>12</sup> Huang
  - <sup>13</sup> Vuksic
  - <sup>14</sup> Delgado
  - <sup>15</sup> Zhonghua, and Limei
  - <sup>16</sup> Folorunso, and Ogunde
  - <sup>17</sup> Tan
  - <sup>18</sup> Tiwari
  - <sup>19</sup> Wegener, and Rüping
  - <sup>20</sup> Gomez-Perez
  - <sup>21</sup> Borrego and Barba
  - <sup>22</sup> Damij and Damij
  - <sup>23</sup> Koh and Low
  - <sup>24</sup> Lee and Siau
  - <sup>25</sup> CRISP (Cross-Industry-Standard-Process)
  - <sup>26</sup> D'heygere
  - <sup>27</sup> Salappa
  - <sup>28</sup> Nissen
  - <sup>29</sup> Ćurko
  - <sup>30</sup> Grigori
  - <sup>31</sup> Mathew, and George
  - <sup>32</sup> Ghanadbashi
  - <sup>33</sup> Gröger
  - <sup>34</sup> Pivk
  - <sup>35</sup> Rupnik and Jaklic
  - <sup>36</sup> Sohail and Dhanapal Durai Dominic
  - <sup>37</sup> Ghattas
  - <sup>38</sup> Claes and Poels
  - <sup>39</sup> Dumas
  - <sup>40</sup> Song
  - <sup>41</sup> Castellanos
  - <sup>42</sup> Yang
  - <sup>43</sup> Davies-Bouldin
  - <sup>44</sup> Filter
  - <sup>45</sup> Wrapper
  - <sup>46</sup> Embedded
  - <sup>47</sup> coefficient of variation
  - <sup>48</sup> standard deviation