

Study of Various Data Mining Methods to Select the Appropriate Method for Managers to Make Decisions in Urban Management (Case Study: Tehran Municipality)¹

Shayesteh Shojaei Karizaki

PhD. Student, Department of Information Science and Knowledge, Information Retrieval, Tonekabon Branch, Islamic Azad University, Mazandaran, Iran. Shojae.sh776@gmail.com

Soodabeh Shapoori

Assistant Professor, Department of Information Science and Knowledge, Information Retrieval Orientation, Tonekabon Branch, Islamic Azad University, Mazandaran, Iran (**Corresponding Author**). shapoori110@yahoo.com

Hajar Zarei

Assistant Professor, Department of Information Science and Knowledge, Information Retrieval Orientation, Tonekabon Branch, Islamic Azad University, Mazandaran, Iran. hajar_zarei@yahoo.com

Abstract

Purpose: The main purpose of this article is to analyze the data of the Tehran Municipality websites and provide data mining solutions for managers' decisions.

Methodology: This research is fundamental and in terms of nature, it can be considered analytical. The data collection method was the field. The statistical population was selected from 220 domains of Tehran Municipality and for analysis, data mining techniques were used to discover the appropriate decision model of city managers. The source of data collection was web analytics and tools used by Google Analytics.

Findings: The accuracy of the LSTM deep neural network is 99.84%. Network accuracy is 99.90%, the call is equal to 99.63%, the error is equal to 0.16%, MSM standard is equal to 0.003. The accuracy of the DBScan method with other basic methods for analyzing the data of Tehran Municipality websites is 99.84%, the deep learning method is 99.25%, the nearest neighbor method is 99.81% and the decision tree method is equal to 99.8%.

1. **The present study is taken from:** Shayesteh Shojaei Karizaki's doctoral dissertation entitled: **Identifying and analyzing data of Tehran Municipality websites and presenting data mining solutions for decision making of its managers**, The supervisor is Soodabeh Shapoori, Department of Information Science and Knowledge, Information Retrieval, Islamic Azad University, Tonekabon Branch, presented on 2020/12/05.

Received: 2020/11/26 ; **Accepted:** 2021/01/27

Shojaei Karizaki, Shayesteh; Shapoori, Soodabeh & Zarei, Hajar (2021). Study of Various Data Mining Methods to Select the Appropriate Method for Managers to Make Decisions in Urban Management (Case Study: Tehran Municipality). *Sciences and Techniques of Information Management*, 7(4).

DOI: 10.22091/stim.2021.6277.1487

© the authors / **Publisher:** University of Qom



With these interpretations, the rate of improvement of the accuracy of the DBScan method in comparison with the deep learning methods is equal to 0.59%.

Conclusion: Finally, by simulating the DBScan method to identify and analyze the data of Tehran Municipality websites and provide data mining solutions for managers' decisions, it was observed that the proposed method provides suggestions to managers to improve site visits and The performance of the municipality is significantly effective.

Keyword: Tehran Municipality, Data Mining, LSTM Deep Nervous Network, Clustering, DBScan.

بررسی روش‌های داده‌کاوی در تصمیم‌گیری مدیران (مطالعه موردی: شهرداری تهران)^۱

شایسته شجاعی کاریزکی

دانشجوی دکتری، گروه علم اطلاعات و دانش‌شناسی، گرایش بازیابی اطلاعات، واحد تنکابن، دانشگاه آزاد اسلامی، مازندران، ایران. Shojaee.sh776@gmail.com

سودابه شاپوری

استادیار، گروه علم اطلاعات و دانش‌شناسی، گرایش بازیابی اطلاعات، واحد تنکابن، دانشگاه آزاد اسلامی، مازندران، ایران (نویسنده مسئول). shapoori110@yahoo.com

هاجر زارعی

استادیار، گروه علم اطلاعات و دانش‌شناسی، گرایش بازیابی اطلاعات، واحد تنکابن، دانشگاه آزاد اسلامی، مازندران، ایران. hajar_zarei@yahoo.com

چکیده

هدف: هدف این پژوهش تجزیه و تحلیل داده‌های وبسایت شهرداری تهران و ارائه راهکارهای داده‌کاوی برای تصمیم‌گیری مدیران است.

روش‌شناسی: پژوهش حاضر بنیادی بوده و از لحاظ ماهیت تحلیلی است. روش گردآوری داده‌ها به صورت میدانی بوده و جامعه آماری از ۲۲۰ دامنه شهرداری تهران انتخاب شده و برای تجزیه و تحلیل داده‌ها از تکنیک‌های داده‌کاوی برای کشف الگوی مناسب تصمیم‌گیری مدیران شهری استفاده گردید و منبع گردآوری داده‌ها وبسجی و ابزار مورد استفاده گوگل آنالیتیکس است.

یافته‌ها: میزان دقت شبکه عصبی عمیق ال.اس.تی.ام برابر با ۹۹/۸۴٪ است. صحت شبکه برابر با ۹۹/۹۰٪، فراخوانی برابر با ۹۹/۶۳٪، خطای برابر با ۰/۱۶٪، معیار ام.اس.ای ام برابر با ۰/۰۰۳٪ است. دقت روش دی.بی.اسکن با سایر روش‌های پایه جهت تجزیه و تحلیل داده‌های وبسایت‌های شهرداری تهران برابر با ۹۹/۸۴٪، روش یادگیری عمیق برابر با ۹۹/۲۵٪، روش نزدیک‌ترین همسایه برابر با ۹۹/۸۱٪ و روش درخت تصمیم برابر با ۹۹/۸٪ است. با این تفاسیر میزان بهبود دقت روش دی.بی.اسکن در مقایسه با روش‌های یادگیری عمیق برابر با ۰/۵۹٪ است.

نتیجه‌گیری: با شبیه‌سازی روش دی.بی.اسکن جهت شناسایی و تجزیه و تحلیل داده‌های وبسایت‌های شهرداری تهران و ارائه راهکارهای داده‌کاوی برای تصمیم‌گیری مدیران مشاهده گردید که روش مطرح شده، پیشنهادهایی را به مدیران ارائه می‌کند که در جهت بهبود بازدید از سایت و عملکرد شهرداری به میزان قابل توجهی موثر است.

کلیدواژه‌ها: شهرداری تهران، داده‌کاوی، شبکه عصبی عمیق ال.اس.تی.ام، خوشه‌بندی، دی.بی.اسکن، مدیران، وبسایت.

۱. پژوهش حاضر برگرفته از: رساله دکتری شایسته شجاعی کاریزکی با عنوان: **شناسایی و تجزیه و تحلیل داده‌های وبسایت‌های شهرداری تهران و ارائه راهکارهای داده‌کاوی برای تصمیم‌گیری مدیران آن**، استاد راهنما سودابه شاپوری، گروه علم اطلاعات و دانش‌شناسی، گرایش بازیابی اطلاعات، دانشگاه آزاد اسلامی واحد تنکابن، ارائه شده در تاریخ ۱۳۹۹/۰۹/۱۵ است.

تاریخ دریافت: ۱۳۹۹/۰۹/۰۶ ؛ تاریخ پذیرش: ۱۳۹۹/۱۱/۰۸

۱. مقدمه

کشف دانش و داده‌کاوی یک حوزه جدید میان‌رشته‌ای و در حال رشد است که حوزه‌های مختلفی همچون پایگاه داده، آمار، یادگیری ماشین و سایر زمینه‌های مرتبط را با هم تلفیق کرده تا اطلاعات و دانش ارزشمند نهفته در حجم بزرگی از داده‌ها را استخراج نماید. با رشد سریع کامپیوتر و استفاده از آن در دو دهه‌ی اخیر تا حدودی همه سازمان‌ها حجم عظیمی از داده را در پایگاه‌های داده خود ذخیره کرده‌اند و سازمان‌ها به فهم این داده‌ها و یا کشف دانش مفید از آن‌ها نیاز دارند. چنان‌چه هر سازمان به ویژه سازمان‌های شهری از داده‌های خوب، الگوی مناسبی را جهت تصمیم‌گیری و تصمیم‌یاری مدیران خود ارائه دهند، زمینه ایجاد شهر هوشمند فراهم خواهد شد. استفاده از داده‌کاوی در جهت کشف الگوی مناسب تصمیم‌گیری و تصمیم‌یاری مدیران شهری و پیش‌بینی آن از پرکاربردترین موضوعات روز دنیا است. از آنجایی که داده‌کاوی بیشتر به کاربردها و اهداف موردنظر محقق و یا مدیر توجه دارد تا به اهمیت داده‌ها و اطلاعات، تشریح کاربردها و به‌کارگیری آن در تصمیم‌سازی خدمات انرژی مفید و ضروری است. سازمان‌های بزرگ، داده‌های بسیار زیادی را درباره مشتریان، تأمین‌کنندگان و شرکای تجاری خود جمع‌آوری و ذخیره می‌کنند، ولی ناتوانی این سازمان‌ها برای کشف دانش پنهان درون این داده‌ها سبب می‌شود که این داده‌ها به دانش تبدیل نشوند و بدون استفاده بمانند. صاحبان کسب‌وکار میل به استخراج اطلاعاتی ناشناخته، معتبر و قابل درک از بانک‌های اطلاعاتی عظیم خود و استفاده از این اطلاعات برای کسب سود بیشتر دارند (دشپاندی^۱ و ثاکاری^۲، ۲۰۱۰) فرآیند تصمیم‌گیری، اساس مدیریت را در سازمان‌ها تشکیل می‌دهد و اطلاعات صحیح را می‌توان از ملزومات اصلی اخذ تصمیم‌های صحیح در سازمان‌ها به شمار آورد. از سوی دیگر، داده‌کاوی به عنوان یکی از دستاوردهای موثر و مفید فن‌آوری اطلاعات در حوزه استفاده بهینه از داده‌ها، توانسته نقشی اساسی برای دستیابی به اطلاعات مورد نیاز مدیران در سازمان‌ها و نهادها ایفا نماید. تهران کلان‌شهری با داده‌های انبوه است که به نظر می‌رسد در مدیریت شهری آن، تاکنون استفاده بهینه‌ای از داده‌های موجود نشده است. این امر ضرورتی را برای پژوهش حاضر در زمینه شناسایی و معرفی این داده‌ها و استفاده از تکنیک‌ها و فنون مناسب داده‌کاوی، ایجاد نموده است. با رشد فزاینده‌ی جمعیت در

کلان‌شهر تهران و به دنبال آن افزایش مشکلات و مسائل ناشی از شهرنشینی در ابعاد وسیع اقتصادی، اجتماعی، فرهنگی و زیست‌محیطی، لزوم ایجاد و استفاده از بسترهای مدیریت شهری جدید همراه با بروزترین و کاربردی‌ترین فن‌آوری‌ها احساس می‌شود (امیری و همکاران، ۱۳۸۸، ص ۳). تهران با جمعیت بیش از ۱۳ میلیون نفر و مساحت ۷۳۰ کیلومتر مربع، پرجمعیت‌ترین کلان‌شهر ایران محسوب می‌شود. ناگفته پیداست که ارائه خدمات شهری در چنین کلان‌شهری با تراکم بیش از ۱۳/۲۶۷ میلیون نفر بر کیلومتر مربع که سفرهای آونگی^۱ ساکنین شهرهای دور و نزدیک اطراف نیز به جمعیت روز آن می‌افزاید، نیازمند به‌کارگیری تمهیدات ویژه بوده و بدون حضور و نقش فعال شهروندان امکان‌پذیر نمی‌باشد.

در پژوهش‌های پیشین از داده‌کاوی برای تجزیه و تحلیل داده‌های موجود در پایگاه داده شهرداری‌ها بدین صورت استفاده شده که چگونگی استفاده از روش‌های داده‌کاوی برای بهبود عملکرد دوربین‌های کنترل‌کننده چراغ‌های راهنمایی تقاطع‌ها را بررسی کردند، استفاده از این روش‌های هوشمند ایمنی ترافیک را افزایش و تلفات ناشی از تصادفات را کاهش می‌دهد. به این منظور الگوریتم‌های درخت تصمیم سی.ای.آر.تی^۲، شبکه عصبی مصنوعی و الگوریتم خوشه‌بندی کی-میانگین مورد استفاده قرار گرفت. درخت تصمیم مجموعه قوانینی را می‌سازد که در تصمیم‌گیری‌های آینده می‌تواند مؤثر باشد. شبکه عصبی هم با بررسی و تجزیه و تحلیل داده‌های گذشته، اتفاقات آینده را پیش‌بینی می‌کند. مدل خوشه‌بندی کی-میانگین هم تأثیر خوشه‌های همگن بر متغیرهای هدف را وزن‌دهی می‌کند. همه این مدل‌ها با استفاده از داده‌های واقعی جمع‌آوری شده توسط بخش حمل و نقل شهرداری شهرهای مریلند و واشنگتن از تصادفات کشنده بین سال‌های ۲۰۰۰ تا ۲۰۰۳، آموزش داده شده‌اند. استفاده از این روش‌های داده‌کاوی، روابط قابل پیش‌بینی بین داده‌های دموگرافیک راننده‌ها و تصادفات مهلک، نوع تصادفات کشنده، رابطه بین روز و زمان و تصادفات کشنده را شناسایی می‌کند (سولمن^۳ و همکاران، ۲۰۰۶). در کشور ترکیه در تحقیقی بهینه‌سازی تصمیم‌گیری در شهرداری‌ها با استفاده از داده‌کاوی را مورد بررسی قرار دادند. در این تحقیق تلاش شده است تا با استفاده از اطلاعات و داده‌های مربوط به

1. Commute
2. CART
3. Solomon

گذشته، رویکردی هوشمند برای ارائه خدمات در فعالیت‌های آینده، فراهم شود. هدف اصلی این پژوهش اتخاذ تصمیمات هوشمند در شهرداری‌ها و پیش‌بینی فعالیت‌های آینده با استفاده از داده‌های مربوط به گذشته است. این تصمیم‌ها شامل: تجزیه و تحلیل‌های فرهنگی-اجتماعی، تجزیه و تحلیل درآمد بر هزینه، تجزیه و تحلیل زیرساخت‌ها، کشف و شناسایی تقلب^۱، ساده‌سازی، تایید و تجزیه و تحلیل تشابه است. اهداف این پژوهش عبارتند از: فراهم آوردن اطلاعات وضعیت فعلی، تسهیل تصمیم‌گیری در مورد فعالیت‌های آتی، افزایش درآمد و کاهش هزینه‌ها و پشتیبانی ساده سیستم ردیابی اسناد. سیستم پیشنهادی در این پژوهش بر مبنای معماری سرویس‌گرا^۲ طراحی شده است. در این سیستم روش‌های داده‌کاوی به صورت یکپارچه با سیستم مدیریت اطلاعات ا.ا.ال.جی.یو (ا.ا.ام.ای.اس)^۳ مورد استفاده قرار گرفته‌اند، بر همین اساس سیستم مذکور ا.ا.ام.آی.اس-دی.ام^۴ نامیده می‌شود. در این سیستم ابتدا ۱۷ سناریو مطرح شده است که شامل: تجزیه و تحلیل کارکنان، دسته‌بندی شهروندان براساس مالیات املاک و مستغلات، توزیع تأخیر مالیات بر املاک و مستغلات شهروندان، تجزیه و تحلیل عملکرد درآمد، تجزیه و تحلیل سوخت، تجزیه و تحلیل مصرف برق، تجزیه و تحلیل صندوق پول، توزیع بنیادهای شرکتی، تجزیه و تحلیل مواد در حرکت، تجزیه و تحلیل ثبت وقایع^۵، تجزیه و تحلیل حساب‌های کاربری، تجزیه و تحلیل حسابداری، تجزیه و تحلیل کارمندان، برآورد دستمزدها، آنالیز شهروندان و تجزیه و تحلیل بنیادهای شرکتی است. در ادامه برای این ۱۷ جریان ۵ مدل: خدمات وب با استفاده از قوانین انجمنی^۶، خدمات وب با استفاده از تجزیه و تحلیل تشخیص نقاط پرت^۷، خدمات وب با استفاده از دسته‌بندی سی.دابلیو.اس^۸، خدمات وب با استفاده از خوشه‌بندی کلوست دابلو.اس^۹، خدمات وب با استفاده از آماده‌سازی داده^{۱۰} توسعه داده شده است. از بین مدل‌های توسعه

1. Fraud
2. Service Oriented Architecture (SOA)
3. OLGU (OMIS)
4. OLGU Management Information System-Data Mining (OMIS-DM)
5. Logs Analyzing
6. Association Rule Mining Web Service (ARMWS)
7. Outlier Detection Analysis Web Service (ODAWS)
8. Classification Web Service
9. Clustering Web Service
10. Data Preparation Web Service

داده شده در ۷ جریان از ای.آر.ام.دبلیو.اس در ۳ جریان از اُدی.ای.دبلیو.اس، در ۲ جریان از سی.دبلیو.اس و ۵ جریان از کلوست دبلیو.اس استفاده شده است (بکیرلی^۱ و همکاران، ۲۰۱۲). از روش‌های داده‌کاوی برای مدیریت مؤثر روابط با شهروندان استفاده کرده‌اند. در این پژوهش یک مدل داده‌کاوی ترکیبی برای ایجاد ارتباط مؤثر شهروندان با سازمان شهرداری با توجه به خدمات شهری توسعه پیدا کرده است. در این پژوهش شکایت‌های شهروندان شهر ویساکاپاتنام که توسط شهرداری این شهر جمع‌آوری شده، به عنوان مجموعه داده مورد استفاده قرار گرفته است. در این مدل ترکیبی ابتدا از روش خوشه‌بندی الگوریتم کی- میانگین برای شناسایی شکایاتی که فرکانس بالایی دارند، استفاده شده است، سپس از روش قوانین انجمنی (الگوریتم آپریوری) برای تعیین عواملی که روی نرخ رضایت در این شکایت‌ها تأثیر دارند، استفاده شده است (لاتها^۲ و همکاران، ۲۰۱۳). در پژوهشی برای دسته‌بندی مسیر تغییر محلات از روش‌های داده‌کاوی استفاده نمودند. شناسایی فرآیند تغییر محلات در طول زمان برای مسئولین شهرداری و سیاست‌گذاران از اهمیت بالایی برخوردار است، این اطلاعات در تدوین سیاست‌ها مؤثر بوده و چشم‌اندازهای اقتصادی را بهبود می‌بخشد. تحقیقات تجربی در مورد تغییر چهره اجتماعی- اقتصادی محلات به دلیل حجم بالای پیچیدگی‌های فهم تغییرات در طول زمان و عوامل تأثیرگذار زیادی که وجود دارد، بسیار سخت و غیرممکن است، در حالی که استفاده از روش‌های داده‌کاوی، این پیچیدگی‌ها را هموار می‌کند. در این پژوهش سعی شده است تا محلاتی که با توجه به ابعاد و ویژگی‌های چندگانه، تغییر مسیر مشابهی داشتند، در یک دسته قرار گیرند. به این منظور یک روش دو لایه در این پژوهش پیشنهاد شده است، که با نمایش داده‌های زمانی و چندبعدی، در فضای ورودی دو بعدی شروع می‌شود، برای اینکار از الگوریتم نقشه خود سازمان‌دهنده استفاده شده است. سپس مکان هر کدام از محلات روی این فضای ورودی ردیابی می‌شود تا مسیرهای زمانی تغییرات محلات شناسایی شود و در نهایت هم الگوریتم خوشه‌بندی کی- میانگین^۳ برای گروه‌بندی محلاتی که مسیرهای تغییر مشابه دارند، مورد استفاده قرار گرفته است. در این مطالعه از داده‌های جغرافیایی، تاریخی و اقتصادی ۸ شهر مجزا در ایالات متحده استفاده شده که شامل شهرهای شارلوت^۴

1. Bakirli
2. Latha
3. Avg, mod
4. Charlotte

دنور^۱، فونیکس^۲، پورتلند^۳، بوفالو^۴، دیترویت^۵، پیتسبورگ^۶ و شیکاگو^۷ است. در این مطالعه از پایگاه داده‌ای که شامل اطلاعات داده‌های سرشماری سالانه از سال‌های ۱۹۷۰ تا ۲۰۱۰ است، استفاده شد. داده‌ی نهایی مورد استفاده دارای ۱۲ ویژگی در سه دسته‌ی اقتصادی-اجتماعی، مسکن و جمعیت قرار می‌گیرند (لینگ^۸ و همکاران، ۲۰۱۶).

در امارات، برای پیش‌بینی تصمیم‌گیری‌ها در ارائه خدمات شهری، از روش‌های داده‌کاوی استفاده کرده‌اند. رشد سالانه جمعیت در مناطق خاصی از شهر رخ می‌دهد و باعث رشد نیاز به خدمات در آن منطقه می‌شود، این نیازها شامل نیاز به مدارس جدید، بیمارستان‌ها، امکانات عمومی، امکانات جاده‌ای، پارک‌های عمومی و... است. مقامات شهری باید به این نیازها رسیدگی کنند. در این پژوهش برای ارزیابی نیازها به منظور ساده‌سازی فرآیند تصمیم‌گیری، مدل‌هایی ارائه شده که نسبت تقاضا بر نیازها را ارزیابی می‌کنند تا مشخص کنند که آیا به برطرف کردن تقاضای مذکور پرداخته شود یا خیر و دلایل رد آن نیز بیان می‌شود. به این منظور در این پژوهش از ۴ روش دسته‌بندی درخت تصمیم، دسته‌بندی بیزی ساده، ماشین‌های بردار پشتیبان و روش دسته‌بندی کی نزدیک‌ترین همسایه استفاده شده است. داده مورد استفاده در این پژوهش از دو منبع تهیه شده است. نخست داده‌های مربوط به نیازهای شهروندان بوده که از شهرداری شهر العین امارات تهیه شده است، که شامل اطلاعات مربوط به نیازهای شهروندان بین سال‌های ۲۰۱۰ تا ۲۰۱۴ می‌باشد. این نیازها از طریق نظرسنجی، کنفرانس، مصاحبه و تشکیل جلسه با شهروندان العین جمع‌آوری شده است. پایگاه داده بعدی مورد استفاده، پایگاه داده مربوط به آمار جمعیت العین در سال ۲۰۱۰ می‌باشد. پایگاه داده نهایی مورد استفاده، که از این دو پایگاه داده یکپارچه‌سازی شده، شامل ۱۲ ویژگی و ۱۰۹۱ رکورد است. این متغیرها شامل: نیازهای تفصیلی، بخش، منطقه، جمعیت کل اماراتی‌ها، جمعیت کل غیر اماراتی‌ها، جمعیت کل، دپارتمان مسئول پیاده‌سازی،

1. Denver
2. Phoenix
3. Portland
4. Buffalo
5. Detroit
6. Pittsburgh
7. Chicago
8. Ling

منبع نیازها و... است.

در این پژوهش عملیات دسته‌بندی، دو بار انجام می‌شود. در دسته‌بندی اول، متغیر هدف پرداختن یا عدم پرداختن به نیاز در نظر گرفته شده است. در دسته‌بندی دوم در صورتی که گزینه‌ی عدم پرداختن به نیاز در دسته‌بندی اول انتخاب شود، یک بار دیگر دسته‌بندی برای بررسی علت عدم پرداختن به نیاز مذکور به صورت یک مسأله دسته‌بندی ۱۲ کلاسه انجام می‌شود (ال نوائمی^۱ و همکاران، ۲۰۱۶). در کشور هند، یک سیستم پشتیبانی تصمیم برای مدیریت پسماندهای جامد شهری ارائه کردند. امروزه بسیاری از کشورها با مسأله مدیریت پسماندهای جامد شهری روبه‌رو هستند. با افزایش نرخ رشد جمعیت، حجم زباله‌های تولیدی نیز افزایش یافته است. در حال حاضر تمرکز زیادی بر تمام سطوح تفکیک زباله مانند روند تولید زباله و تکنولوژی‌های موجود برای جمع‌آوری و پردازش آن وجود دارد. تکنیک‌های تحلیلی جدید بسیاری برای مقابله با مسائل مربوط به مدیریت پسماندهای شهری ارائه شده است. این تکنیک‌ها در تجزیه و تحلیل داده‌های خام پسماندهای جامد شهری کمک زیادی به سازمان‌ها کرده است، که روش‌های داده‌کاوی یکی از این تکنیک‌ها به شمار می‌آید. این روش‌های تحلیلی به تصمیم‌گیران کمک می‌کند تا تصمیم‌های صحیحی اتخاذ کرده و یا به بهبود کارایی تصمیم‌گیری کمک می‌کنند. یک چارچوب عالی برای مدیریت پسماند جامد باید تمام عوامل مختل‌کننده مانند تولید آلودگی، استفاده از زمین، استفاده از انرژی، هزینه‌های مالی، نیازهای نیروی کار، نظارت، نرخ بازیافت و... را شامل شود. در سیستم‌های پشتیبان تصمیم، برای یک فرآیند تصمیم‌گیری دقیق، مدیران به دانش نیاز دارند. وقتی حجم داده‌ها افزایش پیدا می‌کند، مشکلاتی در تجزیه و تحلیل داده‌ها برای ایجاد الگوهای دانش پیش می‌آید، در این‌گونه موارد استفاده از روش‌های داده‌کاوی در کشف الگوهای دانش، انسان را یاری می‌دهد. در این پژوهش از روش خوشه‌بندی کی- میانگین برای ایجاد سیستم پشتیبان تصمیم استفاده گردیده است. در این پژوهش از دو مجموعه داده استفاده شده و یک مجموعه داده مربوط به مناطق و یک مجموعه داده مربوط به مدیریت پسماندهای جامد بوده که توسط شهرداری جمع‌آوری شده است (شارما^۲ و همکاران، ۲۰۲۰).

از روش‌های داده‌کاوی در مدیریت ارتباط با شهروندان استفاده نمودند. در مدیریت ارتباط با شهروند (سی.آی.آر.ام^۱) دولت الکترونیک در جوامع امروزی در حال گسترش است. جامعه ایران نیز از این امر مستثنی نیست. بنابراین، لزوم توجه به مفاهیمی که در این زمینه مطرح می‌شود، امری ضروری است. مدیریت ارتباط با شهروند یکی از این موارد است که بر شهروند محوری تمرکز دارد. در مدیریت ارتباط با شهروند، هدف اصلی شناخت نیازها و خواسته‌های شهروندان و تلاش در جهت بهبود ارائه خدمات و افزایش سطح رضایت‌مندی ایشان است. شهرداری به عنوان یکی از کلیدی‌ترین سازمان‌های دخیل در این امر، نقش عمده‌ای در ارائه خدمات شهری به شهروندان ایفا می‌کند. این سازمان، سامانه‌ای با عنوان سیستم تحلیل اطلاعات ۱۳۷ ایجاد نموده که مرکز ساماندهی و رسیدگی به معضلات و مشکلات شهری است و پل ارتباطی بین شهروندان و شهرداری محسوب می‌شود. اطلاعات حاصل از این سامانه، حاوی مطالب مفیدی در مورد خدمات ارائه شده به شهروندان است و می‌تواند به عنوان منبعی مهم و مناسب برای مدیریت روابط تحلیلی با شهروند مورد بررسی و کاوش قرار بگیرد. از این‌رو با استفاده از قوانین انجمنی تماس‌های شهروندان را مورد تجزیه و تحلیل قرار داده و ارتباط بین مشکلات، مناطق و نواحی آن‌ها مشخص شده است. همچنین با در نظر گرفتن شرایط مکانی، دنباله بروز مشکلات نیز به دست آمده است. انتظار می‌رود نتایج به دست آمده در مدیریت مشکلات شهری و افزایش سطح رضایت‌مندی شهروندان مثرتر واقع شود (مینایی بیدگلی و همکاران، ۱۳۸۸). استفاده از تکنیک داده‌کاوی برای پایش، پیش‌بینی و مدیریت اجرای اثربخش طرح تفصیلی (نمونه موردی شهرداری منطقه ۱۱ مشهد)، طرح‌های جامع و توسعه شهری اسناد بالادستی الگویی مشخص برای مدیریت ظرفیت و قابلیت‌های توسعه یک شهر را مشخص می‌کنند. علاوه بر این بازنگری طرح تفصیلی شهر مشهد و همچنین طرح جامع شهر مشهد و وجود طرح‌های توسعه شهری مختلف از مواردی است که ساخت و سازها را شامل می‌شود. افزایش جمعیت، جابجایی جمعیت در شهرها، رشد تقاضا برای مسکن و همچنین افزایش درآمد سرانه مردم در سال‌های اخیر و نیز افزایش گردش مالی در حوزه اقتصاد مسکن از مواردی است که رشد و توسعه شهری را دچار تغییر و تحولات زیادی نموده است. رفتار مردم در ساخت و سازها آن‌جا که برای أخذ پروانه احداث

ساختمانی به شهرداری مراجعه می‌نمایند و پس از دریافت آن، یکی از مولفه‌های مهم در شناسایی نیازمندی‌های شهروندان بوده و می‌تواند مبنایی برای برنامه‌ریزی‌های توسعه شهری برای برنامه‌ریزان شهری باشد.

پژوهشی با استفاده از تکنیک داده‌کاوی و امکاناتی که استفاده از این ابزار در اختیار تصمیم‌سازان و تصمیم‌گیران مدیریت شهری می‌گذارد، به بررسی، تحلیل و پیش‌بینی رفتار مردم در ساخت و سازهای شهری (نمونه موردی منطقه ۱۱ شهرداری مشهد) براساس پروانه‌های ساختمانی صادر شده پرداخته و براساس آن هنجارها و ناهنجاری‌های رفتاری ساخت‌وساز شهری را شناسایی نموده و هم‌چنین نسبت به معرفی سامانه‌ای که به بررسی، کنترل، پایش و پیش‌بینی مدل‌ها و الگوهای رفتاری مردم می‌پردازد، براساس تکنیک‌های داده‌کاوی اقدام می‌کند، اقدام شده است (خوش‌قامت و همکاران، ۱۳۹۳).

پژوهش دیگری با طرح این مسأله که هدایت‌پذیری دانش در بین گروه‌های مختلف دانشی، به سختی انجام می‌شود، به ارائه الگویی برای حکمرانی دانش در این مورد با استفاده از شبکه‌های اجتماعی پرداخته است. از آنجا که مرور ادبیات، به خوبی این باور را که حکمرانی دانش تنها در حد مطالعات کتابخانه‌ای و نه میدانی، انجام شده است را به اثبات می‌رساند و نشانه‌های لازم برای ارائه یک طراحی منسجم را نیز ارائه نمی‌نماید. از این‌رو محقق، با فهم وضعیت، نسبت به استفاده از روش تحقیقات علم طراحی به عنوان روشی اقدام محور، درصدد فهم وضعیت، شناسایی نیازها و اهداف طراحی، ارائه طراحی و سپس ایجاد مصنوع لازم و در نهایت پیاده‌سازی آن در مورد مطالعه، اقدام نموده است. استفاده از ابزار شبکه‌های اجتماعی به عنوان یک فناوری اجتماعی که می‌تواند سبب تحول‌آفرینی شود، کمک شایانی به ارائه قابلیت‌های مورد نیاز در طراحی منحصربه‌فرد در حوزه حکمرانی دانش نمود. نتایج نشان داد که با استفاده از منطق‌های نوین این قبیل فناوری‌ها می‌توان بسیاری از مشکلاتی که ناشی از ترس، عدم اعتماد، نبود انگیزه، عدم فعالیت و... قلمداد می‌شود را برطرف نمود (قدوسی و همکاران، ۱۳۹۵).

با بررسی و مرور پژوهش‌های انجام شده کاربردهای گوناگون داده‌کاوی در شهرداری‌ها مشخص گردید و پایگاه‌های داده مورد استفاده در این پژوهش‌ها، بررسی شده‌اند. نتیجه مطالعات نشان می‌دهد که شهرداری‌ها مجموعه داده‌های متنوع و مفیدی دارند که در صورت استفاده از آن‌ها و شناسایی الگوها و دانش پنهان در آن‌ها، می‌توانند کمک قابل توجهی در فرآیند مدیریت شهری و

تصمیم‌گیری‌ها داشته باشند و در سیاست‌گذاری‌های شهرداری و در نهایت در بسترسازی ایجاد شهر هوشمند مورد استفاده قرار گیرند. قلمرو این پژوهش از منظر مکانی-زمانی، در حوزه‌ی کاربرد داده‌کاوی در مدیریت سازمان فن‌آوری اطلاعات و ارتباطات در شهر تهران به علاوه دبیرخانه و مرکز تهران هوشمند است. محورهای مأموریتی و حوزه‌های کاری سازمان فاوا^۱ عبارت است از: سامانه‌ها و نرم‌افزارهای مدیریت شهری، سرویس‌های شهروندی و خدمات پرداخت، راهبری شهر هوشمند، داده و آمار و رصد شهری، زیرساخت شبکه برق اضطراری و سخت‌افزار، برنامه‌ریزی و آموزش و پژوهش و خدمات مدیریت، مرکز داده و امنیت. برقراری ارتباطات فی‌مابین کلیه ساختمان‌های شهرداری تهران از طریق فیبرنوری یا وایرلس، راه‌اندازی اتوماسیون آموزش و اتوماسیون اداری به صورت یکپارچه جهت استفاده کلیه واحدهای شهرداری تهران و اجرای پروژه برق اضطراری جهت تمامی ساختمان‌های شهرداری از اقدامات اساسی صورت گرفته بوده که نقش مهمی در پژوهش‌های سابق نداشته است، ولیکن با توجه به تحولات مؤثر و بسیاری که رخ داده، پژوهش در این زمینه ضروری به نظر می‌رسد. بنابراین، هدف پژوهش توصیفی-تحلیلی حاضر بررسی مدل داده‌ای یک پروژه داده‌های معمولی در یک محیط اطلاعاتی کلان برای فرصت‌هایی جهت بهبود نمایش اطلاعات است.

۲. روش‌شناسی پژوهش

از آنجا که در پژوهش حاضر، هدف، توصیف روابط و الگوهای موجود در پایگاه‌های داده سایت‌های شهرداری و یافتن راهی برای کشف الگوی مناسب در آن‌ها بوده و در اصل دست‌کاری و تغییر داده‌ها مورد نیاز نیست؛ بنابراین، یک پژوهش تحلیلی خواهد بود. از نظر هدف نیز این پژوهش بنیادی است، زیرا به دنبال تعیین روابط تجربی میان پدیده‌ها بوده و هدف، تشخیص الگوها و روابط تحلیلی از داده‌های حاصل از حوادث یک محدوده زمانی خاص بوده و با توجه به این‌که در طرح پژوهش حاضر داده‌های مورد استفاده درون پایگاه‌های داده بوده و اینکه براساس مدل‌های خوشه‌بندی و طبقه‌بندی داده‌کاوی و الگوریتم‌های هوش مصنوعی انجام یافته است، این پژوهش را می‌توان از لحاظ ماهیت، تحلیلی در نظر گرفت. منبع گردآوری داده‌ها وب‌سنجی بوده و ابزار مورد استفاده گوگل آنالیتیکس^۲ است.

۱. سازمان فن‌آوری اطلاعات و ارتباطات

2. Google Analytics

پژوهش بدین صورت اجرا گردید که ابتدا دیتاست^۱ کلیدواژه‌های جمع‌آوری شده به سیستم دی.بی.اسکن^۲ وارد شد. سپس داده‌ها مورد پیش‌پردازش قرار گرفت. در مرحله بعد فرآیند خوشه‌بندی دی.بی.اسکن بر روی داده‌ها اعمال شده و یک خوشه‌بندی اولیه بر روی داده‌ها صورت گرفت. از برچسب‌های خوشه به عنوان کلاس برای نمونه‌ها استفاده شد. سپس داده‌های خوشه‌بندی شده به دو بخش آموزشی و آزمایشی تبدیل گردید. داده‌های آموزشی ۷۰٪ و داده‌های آزمایشی ۳۰٪ از کلیدواژه‌ها را تشکیل می‌دهد. در نهایت به ازای تک‌تک داده‌ها، کلیدواژه‌های مطلوب و نامطلوب استخراج گردید. از این‌رو، با توجه به اینکه داده‌های در دسترس در این پژوهش داده‌های مربوط به کلیدواژه‌های سایت‌های مختلف در شهرداری است، هدف نهایی توصیه کلیدواژه‌های پر طرفدار به مدیران سایت‌ها می‌باشد. این امر موجب می‌گردد که کلیدواژه‌های بهینه برای سایت‌ها تعیین و در نهایت بازدهی سامانه و سایت‌ها افزایش پیدا کند. مهم‌ترین قسمت‌های روش گردآوری داده‌ها در این پژوهش در زیربخش‌های مربوطه ارائه خواهد شد.

- خوشه‌بندی اولیه کلیدواژه‌ها با استفاده از الگوریتم دی.بی.اسکن

- طبقه‌بندی کلیدواژه با الگوریتم یادگیری عمیق ال.اس.تی.ام^۳

۲-۱. تفکیک نمونه‌های آموزشی و آزمایشی

گوگل آنالیتیکس یک ابزار تحلیلی وب رایگان بوده که توسط گوگل برای کمک به صاحبان مشاغل و سایت‌ها در تحلیل ترافیک وب‌سایت ارائه شده است.

نحوه‌ی کار گوگل آنالیتیکس بدین صورت است: گوگل آنالیتیکس چندین خط کد رهگیری را در کد وب‌سایت سازمان قرار می‌دهد. این کد فعالیت‌های مختلف کاربران را هنگام بازدید از وب‌سایت به همراه ویژگی‌های (مانند سن، جنسیت، علایق) آن کاربران ثبت می‌کند. پس از خروج کاربر از وب‌سایت، تمام اطلاعات را به سرور گوگل آنالیتیکس ارسال می‌کند. در مرحله بعد، گوگل آنالیتیکس داده‌های جمع‌آوری شده از وب‌سایت را به چند روش، در درجه اول با چهار سطح جمع می‌کند:

1. Data
2. DBScan
3. S.T.M

سطح کاربر^۱ (مربوط به اقدامات هر کاربر)

سطح بازدید^۲ (هر بازدید منحصر به فرد)

سطح بازدید صفحات^۳ (هر بازدید منحصر به فرد از صفحه)

سطح رویداد^۴ (کلیک دکمه، نمایش ویدیو و غیره) (موسوی نطنزی، ۱۳۹۸).

در کاربردهای آماری، تجزیه و تحلیل داده‌ها را می‌توان به آمار توصیفی، تجزیه و تحلیل داده‌های اکتشافی^۵ و تجزیه و تحلیل داده‌های تاییدی^۶ تقسیم کرد. روش ای.دی.ای^۷ بر کشف ویژگی‌های جدید در داده‌ها و سی.دی.ای^۸ بر تایید یا تکذیب فرضیه‌های موجود تمرکز دارد. داده‌کاوی در وب‌سایت‌ها داده‌های اطلاعاتی به عنوان یکی از منابع حیاتی سازمان شناخته می‌شود و بسیاری از سازمان‌ها با اطلاعات و دانش سازمانی خود مانند سایر دارایی‌های ارزشمندشان برخورد می‌کنند. داده اطلاعاتی به اطلاعات خام سازمان اطلاق می‌شود و اطلاعات نیز داده‌های پردازش شده است. همچنین داده‌های پردازش شده پس از طبقه‌بندی و آنالیز به دانش سازمان تبدیل می‌گردند (گوبل^۹ و گرووالد^{۱۰}، ۱۹۹۹). داده‌های در دسترس این سایت‌ها عبارتند از: کلیدواژه، امتیاز، تعداد مشاهده، تعداد کلیک، نام سایت. به طور کلی، جامعه آماری این پژوهش داده‌های استخراج شده از بین ۲۲۰ دامنه منتخب در سال ۱۳۹۸-۱۳۹۷ می‌باشد. این پژوهش فاقد روش نمونه‌گیری است، چرا که تحلیل و پشتیبانی تصمیم در مدیریت، نیازمند مشارکت تمامی داده‌های مرتبط با آن می‌باشد. داده‌های جمع‌آوری شده در قالب فایل اکسل ارائه شده و در نهایت به الگوریتم خوشه‌بندی وارد می‌شود. در مرحله بعد مراحل شبیه‌سازی و اعمال الگوریتم خوشه‌بندی جهت ایجاد برجسب اولیه به داده‌ها ارائه شده است.

1. User level
2. Session level
3. Pageview level
4. Event level
5. Exploratory Data Analysis (EDA)
6. Confirmatory Data Analysis (CDA)
7. EDA
8. CDA
9. Goebel
10. Gruenwald

۳. یافته‌ها

خوشه‌بندی اولیه کلیدواژه‌ها با استفاده از الگوریتم دی.بی.اسکن به این صورت است که پس از اینکه داده‌های پیش‌پردازش شده تهیه شد، این داده به الگوریتم خوشه‌بندی اعمال گردید. ابتدا دیتاست توسط کنترلر اکسل‌خوان به سیستم وارد می‌شود. سپس داده‌های پرت توسط کنترلر جایگزینی مقدار مجهول حذف می‌شوند. در مرحله بعد الگوریتم خوشه‌بندی دی.بی.اسکن داده‌ها را مورد پردازش قرار داده و خوشه‌بندی اولیه می‌شود. در نهایت خروجی خوشه‌بندی توسط کنترلر اکسل‌نویس چاپ می‌شود. در این تحقیق سه خوشه به عنوان تعداد خوشه‌های بهینه جهت گروه‌بندی اولیه انتخاب گردید. در این طرح از شبکه‌های عصبی بازگشتی به عنوان یکی از دسته‌بندی‌ها استفاده می‌شود. اساس کار شبکه‌های عصبی بازگشتی در پردازش متن، ساخت مدل زبان است. مدل زبان مشابهت جملات به یکدیگر برای یک رخداد خاص براساس احتمال را بیان می‌کند. ایده استفاده از شبکه‌های عصبی بازگشتی استفاده از مشخصه‌ها و ویژگی‌ها به صورت متوالی است. یکی دیگر از خصوصیات شبکه‌های عصبی بازگشتی، دارا بودن حافظه است که موجب می‌شود، داده‌های محاسبه شده در هر زمان در آن حافظه قرار بگیرند. میزان بهینه‌سازی در هر گام از این شبکه به وسیله محاسبه مقدار ضرر به دست می‌آید. هم‌چنین از ضرر در گام آخر برای تعیین دسته نیز استفاده می‌شود. شبکه‌های عصبی بازگشتی به انواع مختلفی ارائه شده‌اند، یکی از این نوع شبکه‌ها که در طرح حاضر از آن استفاده شده، شبکه‌های عصبی بازگشتی ال.اس.تی.ام است. بدین صورت که با استفاده از الگوریتم، ابتدا براساس داده‌های آموزشی، مدلی تولید می‌گردد. سپس کلیدواژه‌های جدید به مدل مربوطه داده شده و کلیدواژه موردنظر طبقه‌بندی می‌شود. اگر برچسب (۱) به کلیدواژه تعلق گرفت، بدین معنی است که کلیدواژه موردنظر پرت‌فردار بوده و می‌تواند برای سامانه استفاده گردد. در صورتی که به کلیدواژه مورد نظر برچسب (۰) تعلق گیرد، بدین معنی است که کلیدواژه کم‌طرفدار بوده و در ایندکس سایت و سامانه تاثیر بسزایی ندارد. پس از اینکه داده‌های اولیه خوشه‌بندی شدند، اقدام به تفکیک داده‌های گروه‌بندی شده به دو دسته آموزشی و آزمایشی می‌گردد. داده‌های آموزشی ۷۰٪ از کل داده‌های موجود (مجموعه داده کلیدواژه‌ها) را تشکیل داده و جهت آموزش الگوریتم شبکه عصبی عمیق استفاده می‌گردد. داده‌های آزمایشی نیز ۳۰٪ از داده‌ها را تشکیل داده و به منظور ارزیابی روش دی.بی.اسکن استفاده می‌شود. کلیدواژه‌های جدید به عنوان داده‌های آزمایشی به مدل شبکه عصبی عمیق وارد می‌شود و براساس آن محبوبیت کلیدواژه تعیین و به مدیران ارائه می‌گردد. با این

استراتژی کلیدواژه‌های دقیق و مرتبط‌تر به وبسایت‌های شهرداری پیشنهاد می‌شود. در این پروژه معیارهای زیر محاسبه شده و پس از پایان شبیه‌سازی، در خروجی نمایش داده می‌شوند:

معیار دقت: این معیار میزان دقت طبقه‌بندی را محاسبه می‌کند. در واقع مشخص می‌کند که الگوریتم یادگیری استفاده شده تا چه میزان طبقه‌بندی نمونه‌های تست را درست انجام داده است.

معیار صحت: این معیار نیز که مکمل معیار دقت است، صحت طبقه‌بندی الگوریتم یادگیری (کی.ان.ان) را می‌سنجد.

معیار فراخوانی: این معیار نیز مکمل معیار ارزیابی صحت است که جنبه دیگری از میزان درستی یادگیری عمیق ال.اس.تی.ام را مورد سنجش قرار می‌دهد.

معیار خطای طبقه‌بندی: این معیار میزان خطای یادگیری عمیق ال.اس.تی.ام را جهت طبقه‌بندی دیتاست وبسایت‌های شهرداری تهران مشخص می‌کند.

در این قسمت به مقایسه یافته‌های شبیه‌سازی شبکه عصبی عمیق ال.اس.تی.ام پرداخته می‌شود. لازم به ذکر است که در این فاز، شبیه‌سازی بر روی ویژگی‌های اصلی دیتاست صورت می‌گیرد. در شکل زیر یافته‌های حاصل از شبیه‌سازی شبکه عصبی عمیق ال.اس.تی.ام نشان داده شده است.

PerformanceVector

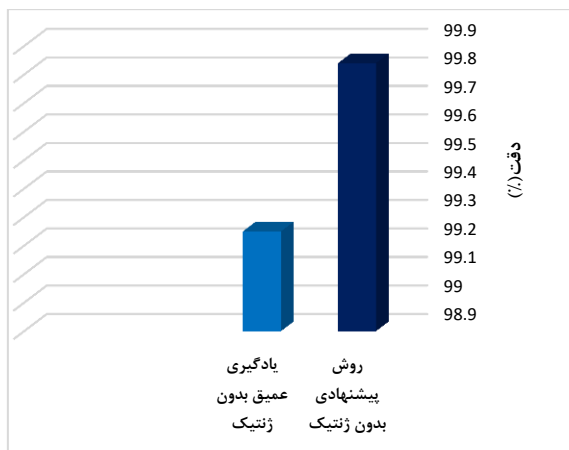
```
PerformanceVector:
accuracy: 99.84% +/- 0.05% (mikro: 99.84%)
ConfusionMatrix:
True:    0.0    1.0
0.0:    5723   12
1.0:     0    1601
classification_error: 0.16% +/- 0.05% (mikro: 0.16%)
ConfusionMatrix:
True:    0.0    1.0
0.0:    5723   12
1.0:     0    1601
weighted_mean_recall: 99.63% +/- 0.12% (mikro: 99.63%), weights: 1, 1
ConfusionMatrix:
True:    0.0    1.0
0.0:    5723   12
1.0:     0    1601
weighted_mean_precision: 99.90% +/- 0.04% (mikro: 99.90%), weights: 1, 1
ConfusionMatrix:
True:    0.0    1.0
0.0:    5723   12
1.0:     0    1601
absolute_error: 0.003 +/- 0.001 (mikro: 0.003 +/- 0.043)
root_mean_squared_error: 0.042 +/- 0.010 (mikro: 0.044 +/- 0.000)
```

شکل ۱- یافته‌های حاصل از شبیه‌سازی شبکه عصبی عمیق ال.اس.تی.ام.

همان‌طور که در شکل (۱) مشاهده می‌گردد، میزان دقت شبکه عصبی عمیق ال.اس.تی.ام برابر با ۹۹/۸۴٪ و صحت شبکه برابر با ۹۹/۹٪، فراخوانی برابر با ۹۹/۶۳٪، خطای برابر با

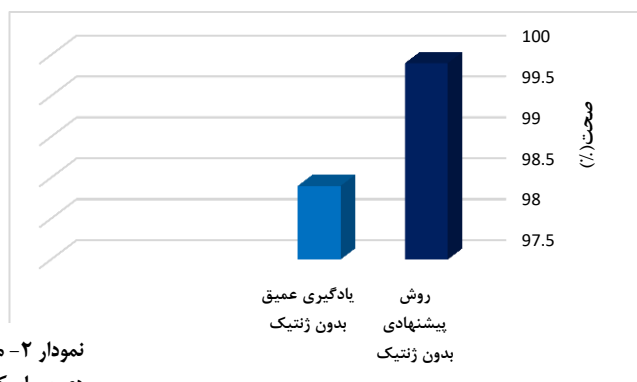
۰/۱۶٪، معیار ام.اس.ای ام. برابر با ۰/۰۰۳ و در نهایت معیار آر.ام.اس.ای. شبکه عصبی عمیق ال.اس.تی. ام برابر با ۰/۰۴۲ است.

در این قسمت به مقایسه روش دی.بی.اسکن با سایر روش‌ها پرداخته می‌شود. در شکل زیر مقایسه دقت روش دی.بی.اسکن با سایر روش‌های پایه جهت تجزیه و تحلیل داده‌های وبسایت‌های شهرداری تهران نشان داده شده است.



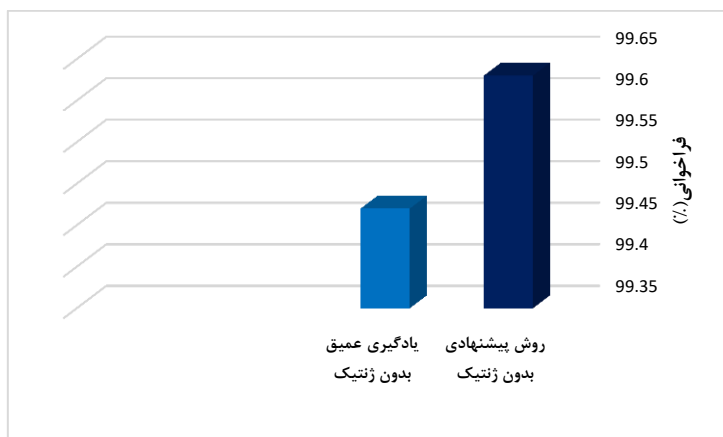
نمودار ۱- مقایسه دقت روش دی.بی.اسکن با یادگیری عمیق

میزان دقت روش دی.بی.اسکن با سایر روش‌های پایه جهت تجزیه و تحلیل داده‌های وبسایت‌های شهرداری تهران برابر با ۹۹/۸۴٪ و روش یادگیری عمیق برابر با ۹۹/۲۵٪ است. با این تفاسیر میزان بهبود دقت روش دی.بی.اسکن در مقایسه با روش‌های یادگیری عمیق برابر با ۰/۵۹٪ است.



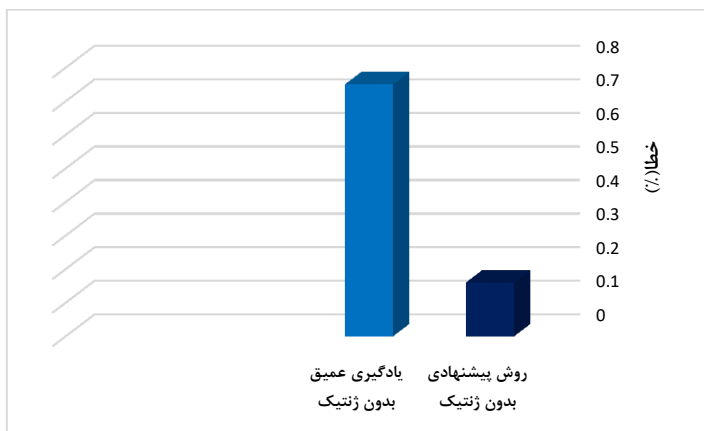
نمودار ۲- مقایسه صحت روش دی.بی.اسکن با یادگیری عمیق

میزان صحت روش دی.بی.اسکن با سایر روش‌های پایه جهت تجزیه و تحلیل داده‌های وبسایت‌های شهرداری تهران برابر با ۹۹/۹٪ و روش یادگیری عمیق برابر با ۹۸/۴٪ است. با این تفاسیر میزان بهبود صحت روش دی.بی.اسکن در مقایسه با روش‌های یادگیری عمیق برابر با ۱/۵٪ است.



نمودار ۳- مقایسه فراخوانی روش دی.بی.اسکن با یادگیری عمیق

میزان فراخوانی روش دی.بی.اسکن با سایر روش‌های پایه جهت تجزیه و تحلیل داده‌های وبسایت‌های شهرداری تهران برابر با ۹۹/۶۳٪ و روش یادگیری عمیق برابر با ۹۸/۴۷٪ است. با این تفاسیر میزان بهبود فراخوانی روش دی.بی.اسکن در مقایسه با روش‌های یادگیری عمیق، نزدیک‌ترین همسایه و درخت تصمیم به ترتیب برابر با ۰/۱۶٪ است.



نمودار ۴- مقایسه خطای روش دی.بی.اسکن با یادگیری عمیق

میزان خطای روش دی.بی.اسکن با سایر روش‌های پایه جهت تجزیه و تحلیل داده‌های وبسایت‌های شهرداری تهران برابر با ۰/۱۶٪ و روش یادگیری عمیق برابر با ۰/۷۵٪ است. با این تفاسیر میزان بهبود خطای روش دی.بی.اسکن در مقایسه با روش‌های یادگیری عمیق برابر با ۰/۵۹٪ می‌باشد.

۴. نتیجه‌گیری

بررسی داده‌های وبسایت‌های شهرداری تهران با این گستره که تمامی کلیدواژه‌ها، کلیک‌ها و بازدیدها را مورد توجه قرار دهد، صرفاً در این پژوهش انجام شده و در پژوهش‌های انجام شده گذشته در خصوص تجزیه و تحلیل داده‌های ثبت شده در سامانه ۱۳۷ که مبنای وجودی آن دریافت دیدگاه‌ها و نقطه نظرات شهروندان در خصوص مدیریت شهری و همچنین ایجاد ارتباط بیشتر شهروندان با مدیران ارشد شهرداری می‌باشد، بوده و انباره عظیم داده‌های وبسایت‌های شهرداری تهران در پایگاه داده‌ای مورد تجزیه و تحلیل و داده‌کاوی برای تصمیم‌گیری و تصمیم‌گیری مدیران تاکنون در هیچ پژوهشی بررسی نشده است. با توجه به اینکه هدف اصلی این پژوهش شناسایی و تجزیه و تحلیل داده‌های وبسایت‌های شهرداری تهران و ارائه راهکارهای داده‌کاوی برای تصمیم‌گیری مدیران بود. بنابراین، در این پژوهش با یک راهکار مبتنی بر شبکه عصبی عمیق و تکنیک خوشه‌بندی اقدام به طبقه‌بندی کلیدواژه‌های مناسب برای مدیران سایت‌های شهرداری شده است. با توجه به یافته‌ها و نتایج بدست آمده مشاهده گردید که دقت، صحت و فراخوانی و سایر یافته‌های تحقیق نسبت به سایر روش‌ها جهت تجزیه و تحلیل داده‌های وبسایت‌های شهرداری تهران و ارائه راهکارهای داده‌کاوی برای تصمیم‌گیری مدیران، بهتر عمل کرده و به این استراتژی که ترکیبی از تکنیک‌های خوشه‌بندی دی.بی.اسکن و شبکه عصبی عمیق ال.اس.تی.ام بوده می‌توان اعتماد کرده و در شهرداری تهران مورد استفاده قرار داد. در نهایت با شبیه‌سازی روش دی.بی.اسکن جهت شناسایی و تجزیه و تحلیل داده‌های وبسایت‌های شهرداری تهران و ارائه راهکارهای داده‌کاوی برای تصمیم‌گیری مدیران مشاهده گردید که روش مطرح شده، پیشنهادهایی را به مدیران ارائه می‌کند که در جهت بهبود بازدید از سایت و عملکرد شهرداری به میزان قابل توجهی موثر است. یافته‌ها نشان داد که وبسایت‌های شهرداری تهران همکاری مطلوبی با یکدیگر نداشته و موازی‌کاری در آن‌ها بسیار است که این خود موجب اتلاف وقت و انرژی و منابع شهرداری‌ها شده و موجب اختلال در اجرای تصمیمات مدیران می‌گردد. از آن‌جا که این

وبسایت‌ها دارای مطالب مشترک در زمینه‌های یکسانی هستند، در صورت تبادل لینک، احتمال دیده شدن آن‌ها افزایش می‌یابد. از آن‌جا که اطلاعات باید جدید و به روز باشد، می‌توان با پشتیبانی و به‌روزرسانی مرتب، امکان مراجعه و تعامل شهروندان با وبسایت‌های شهرداری را افزایش داد و مراجعه فیزیکی و ازدحام ارباب‌رجوع به شهرداری‌ها را کم کرد.

۵. پیشنهادات

- بکارگیری روش‌های ترکیبی، یادگیری تقویتی و یادگیری جمعی به جای الگوریتم شبکه عصبی عمیق ال.اس.تی.ام به منظور ارزیابی نتایج و بهبود یافته‌های تحقیق.
- ترکیب سایر روش‌های یادگیری ماشینی یا تکنیک‌های بهینه‌سازی و انتخاب ویژگی از جمله الگوریتم ژنتیک، الگوریتم گل‌های کلونال^۱ به منظور بهبود تجزیه و تحلیل داده‌های وبسایت‌های شهرداری تهران استفاده از الگوریتم‌های بوستینگ^۲ به منظور بهبود فرایند تجزیه و تحلیل داده‌های وبسایت‌های شهرداری تهران و مقایسه نتایج بدست آمده با یافته‌های این تحقیق.
- ترکیب الگوریتم شبکه عصبی و ماشین بردار پشتیبان با الگوریتم پی.اس.اُ و مقایسه یافته‌های این تحقیق با نتایج بدست آمده با ترکیب روش‌های مطرح شده.

منابع

- امیری، آ.؛ دلجو، غ.؛ قربانی زاده، و. (۱۳۸۸). عوامل مؤثر بر پذیرش سامانه مدیریت شهری تهران (۱۳۷) توسط شهروندان. *مدیریت شهری*، ۶(۲۲): ۷-۲۲.
- خوش قامت، ک.؛ صفدری، ن.؛ یغمایی فر، ش. (۱۳۹۳). *ششمین کنفرانس ملی برنامه ریزی و مدیریت شهری با تأکید بر مؤلفه های شهر اسلامی*. مشهد: شهرداری، شورای اسلامی، دانشگاه فردوسی.
- قدوسی، م.ح.؛ محمودی، ج.؛ موسی خانی، م.؛ جعفری، م. (۱۳۹۵). ارائه چارچوب حکمرانی دانش در شبکه های اجتماعی. *مدیریت علمی اطلاعات*، ۱۵(۱): ۱۵-۳۹.
- موسوی نظنزی، م.ا. (۱۳۹۸). *گوگل آنالیتیکس چیست و چگونه به آنالیز سایت می پردازد؟*. قابل دسترس در: <https://hamyar.co/what-is-google-analytics/>
- مینایی بیدگلی، ب.؛ آخوندزاده، ا.؛ موسوی، م.ح.؛ احمدوند، ع.م. (۱۳۸۸). *استفاده از داده کاوی در مدیریت ارتباط با شهروند: موردکاوی سامانه ۱۳۷ شهرداری تهران*. در: تهران: سومین کنفرانس داده کاوی.

References

- Al Nuaimi, E.; Al Marzooqi, S. & Zaki, N. (2016). **Predicting the decision for the provision of municipal services using data mining approaches**. Paper presented at the 2016 3rd MEC International Conference on Big Data and Smart City (ICBDSC).
- Amiri, A.; Deljoo, Gh.H. & Ghorbani Zadeh, V. (2008). Factors Affecting the Acceptance of Tehran Urban Management System (137) by Citizens. *Journal of Urban Management*, 6(22): 7-22. [In Persian]
- Bakırlı, G.; Birant, D.; Mutlu, E.; Kut, A.; Denktas, L. & Cetin, D. (2012). **Data mining solutions for local municipalities**. Paper presented at the Proceedings of the 12th European Conference on e-Government: ECEG.
- Deshpande, S. & Thakare, V. (2010). Data mining system and applications: A review. *International Journal of Distributed and Parallel systems (IJDPSS)*, 1(1): 32-44.
- Goebel, M. & Gruenwald, L. (1999). A survey of data mining and knowledge discovery software tools. *ACM SIGKDD explorations newsletter*, 1(1): 20-33.
- Khoyqamat, K.; Safdari, N. & Yaghmaefar, Sh. (2015). **Sixth National Conference on Urban Planning and Management with emphasis on the components of the Islamic city**. Mashhad: Mashhad Municipality. Mashhad: Islamic Council, Ferdowsi University of Mashhad. [In Persian]
- Latha, P.P.; Madhuri, R.; Rao, K.P. & Bharathi, M.V. (2013). Effectual citizen relationship management with data mining techniques. *GMR Institute of Technology*, 2: 158-161.
- Ling, C. & Delmelle, E.C. (2016). Classifying multidimensional trajectories of neighbourhood change: a self-organizing map and k-means approach. *Annals of GIS*, 22(3): 173-186.
- Minaei Bidgoli, B.; Akhundzadeh, E.; Mousavi, M.H.; Ahmadvand, A.M. (2009). **Using Data Mining in Citizen Relationship Management: A Case Study of Tehran Municipality 137 System**. In: Electronic City Conference. [In Persian]

- Mousavi Natanzi, M.A. (2020). **What is Google Analytics and how does it analyze the site?**. Available at: <https://hamyar.co/what-is-google-analytics>. [In Persian]
- Qudusi, M.R.; Mahmoudi, J.; Musa Khani, M. & Jafari, M. (2016), Presenting the framework of knowledge governance in social networks. *Bi-Quarterly Journal of Scientific Information Management*, 5(1): 15-39. [In Persian]
- Sharma, N.; Litoriya, R. & Sharma, A. (2020). **Application and Analysis of K-Means Algorithms on a Decision Support Framework for Municipal Solid Waste Management**. Paper presented at the International Conference on Advanced Machine Learning Technologies and Applications.
- Solomon, S.; Nguyen, H.; Liebowitz, J. & Agresti, W. (2006). Using data mining to improve traffic safety programs. *Industrial Management & Data Systems*, 106(5): 621-643.
- DOI:** 10.1108/02635570610666412.

استناد به این مقاله

DOI: 10.22091/stim.2021.6277.1487

شجاعی کاریزکی، شایسته؛ شاپوری، سودابه؛ زارعی، هاجر (۱۴۰۰). بررسی روش‌های داده‌کاوی در تصمیم‌گیری مدیران (مطالعه موردی: شهرداری تهران). *علوم و فنون مدیریت اطلاعات*، ۷(۴): ۹۳-۱۱۴.