



ORIGINAL RESEARCH PAPER

Development of Data-Driven Framework Based on Spatial Analysis and Graph Neural Networks for Service Demand Forecasting and Optimization of Banking Branch Networks

A. Rezaei¹, M. Ghasemi^{1,2}, S.M Tafreshi¹¹ Department of Management, ToH.C., Islamic Azad University, Torbat Heydarieh, Iran² Department of Economics, Management and Accounting, Payame Noor University, Tehran, Iran

ABSTRACT

Received: 07 December 2025
 Reviewed: 19 January 2026
 Revised: 01 May 2026
 Accepted: 15 June 2026

KEYWORDS:

Spatial Data Mining
 Machine Learning
 Smart Banking
 Demand Forecasting
 Model Explainability
 Spatio-Temporal Graph Neural Networks

* Corresponding author

✉ ghasemi.mo@pnu.ac.ir

① (+9851) 38683900

Background and Objectives: The rapid growth of digital banking, electronic payment systems, and online financial services has generated vast amounts of transactional and customer-related data. These developments have created new opportunities for data-driven decision-making in the banking sector, particularly in demand forecasting, service planning, and branch network optimization. Because banking demand is influenced by spatial, demographic, and socioeconomic factors, integrating geospatial analysis with artificial intelligence techniques can provide valuable insights for strategic banking decisions. Despite the increasing availability of banking data, the practical use of spatial data mining and machine learning for demand analysis and branch planning remains limited. Therefore, this study aims to develop and evaluate an integrated framework that combines spatial analysis, machine learning, deep learning, and explainable artificial intelligence to forecast banking service demand and support spatial decision-making for branch network optimization.

Methods: The study was conducted using customer and branch data obtained from Bank Sepah in Markazi Province, Iran, covering the period from 2021 to 2023. Markazi Province was selected as the study area because of its diverse urban, industrial, rural, and service-oriented characteristics, which provide a suitable environment for investigating heterogeneous patterns of banking demand. The proposed framework consists of four main stages. First, transactional, spatial, and socioeconomic datasets were integrated and processed through data cleaning, preprocessing, and feature engineering procedures. Second, spatial demand patterns were investigated using exploratory spatial data analysis, including Moran's I index, Local Indicators of Spatial Association, and density-based spatial clustering methods. Third, predictive models were developed using Random Forest, Extreme Gradient Boosting, Long Short-Term Memory networks, and Spatio-Temporal Graph Neural Networks. Finally, model outputs were interpreted using feature importance analysis and Shapley Additive Explanations to identify the key drivers of banking demand and support decision-making processes. Model performance was evaluated using standard statistical measures, including the coefficient of determination, precision-recall analysis, calibration assessment, and classification performance metrics.

Findings: The results revealed significant spatial heterogeneity in banking service demand across the study area. Spatial analysis produced a Moran's I value of 0.71, indicating strong positive spatial autocorrelation and confirming the existence of clustered demand patterns. The identified hotspots were primarily concentrated in economically active and densely populated counties. Comparative evaluation of predictive models demonstrated that tree-based machine learning algorithms outperformed deep learning approaches for the available dataset. Random Forest achieved the highest predictive performance with an R^2 value of 0.95, followed closely by Extreme Gradient Boosting with an R^2 value of 0.94 for forecasting monthly customer visits. In contrast, the Spatio-Temporal Graph Neural Network model achieved a lower performance level ($R^2 \approx 0.51$). Explainability analysis showed that spatial accessibility, customer socioeconomic characteristics, and historical transaction behavior were among the most influential factors affecting demand patterns. The results further

demonstrated that the integration of spatial features substantially improved predictive accuracy and enhanced the interpretability of the generated models.

Conclusion: The findings indicate that combining spatial analytics, machine learning, and explainable artificial intelligence provides an effective framework for understanding and forecasting banking service demand. Beyond model comparison, the proposed framework supports practical decision-making related to branch network development, service allocation, and digital banking expansion. The identified spatial demand patterns can assist banking managers in prioritizing investment locations, improving service coverage, and optimizing operational resources. However, the limited temporal depth of the available dataset restricted the effectiveness of advanced spatio-temporal deep learning models. Future studies may benefit from longer time-series datasets and real-time transactional information to further enhance predictive performance and spatio-temporal modeling capabilities. Overall, the proposed framework offers a practical and scalable approach for developing intelligent decision-support systems in data-driven banking environments.



NUMBER OF REFERENCES
38



NUMBER OF FIGURES
30



NUMBER OF TABLES
5

مقاله پژوهشی

توسعه یک چارچوب داده‌محور مبتنی بر تحلیل مکانی و یادگیری ماشین برای پیش‌بینی تقاضای خدمات و پشتیبانی از بهینه‌سازی شبکه شعب بانکی

علی رضائی^۱، محمد قاسمی^{۲*}، سید محمد نفرشی^۱

^۱ گروه مدیریت، واحد تربت حیدریه، دانشگاه آزاد اسلامی، تربت حیدریه، ایران

^۲ گروه اقتصاد، مدیریت و حسابداری دانشگاه پیام نور، تهران، ایران

چکیده

پیشینه و اهداف: تحول دیجیتال در نظام بانکی و گسترش بانکداری الکترونیکی، حجم گسترده‌ای از داده‌های تراکنشی، مکانی و اقتصادی-اجتماعی را تولید کرده است که می‌تواند به عنوان منبعی ارزشمند برای تحلیل رفتار مشتریان، پیش‌بینی تقاضای خدمات و پشتیبانی از تصمیم‌گیری‌های راهبردی مورد استفاده قرار گیرند. با این حال، بهره‌گیری نظام‌مند از این داده‌ها در برنامه‌ریزی مکانی خدمات بانکی، مکان‌یابی شعب و تخصیص بهینه منابع در بسیاری از بانک‌ها همچنان محدود است. از سوی دیگر، توزیع تقاضای خدمات بانکی ماهیتی فضایی داشته و تحت تأثیر عوامل جمعیتی، اقتصادی و مکانی قرار می‌گیرد؛ از این رو، تحلیل این الگوها نیازمند رویکردهایی است که بتوانند اطلاعات مکانی را در کنار داده‌های تراکنشی و رفتاری پردازش کنند. هدف پژوهش حاضر، توسعه و ارزیابی یک چارچوب یکپارچه مبتنی بر تحلیل مکانی، یادگیری ماشین و روش‌های تبیین‌پذیر به منظور پیش‌بینی تقاضای خدمات بانکی و پشتیبانی از تصمیم‌گیری‌های مرتبط با توسعه و بهینه‌سازی شبکه شعب است.

روش‌ها: پژوهش حاضر به صورت مطالعه موردی بر داده‌های مشتریان و شعب بانک سپه در استان مرکزی طی سال‌های ۱۴۰۰ تا ۱۴۰۲ انجام شد. استان مرکزی به دلیل برخورداری از ساختار متنوع شهری، صنعتی، روستایی و خدماتی به عنوان منطقه مطالعه انتخاب گردید. چارچوب پیشنهادی شامل مراحل گردآوری و یکپارچه‌سازی داده‌ها، پیش‌پردازش و مهندسی ویژگی‌ها، تحلیل و خوشه‌بندی مکانی تقاضا، مدل‌سازی پیش‌بینانه و تبیین نتایج است. در بخش تحلیل فضایی از شاخص موران‌آی (Moran's I)، تحلیل نقاط داغ، خوشه‌بندی DBSCAN و شاخص‌های خودهمبستگی محلی استفاده شد. در بخش پیش‌بینی، عملکرد الگوریتم‌های LSTM، XGBoost، Random Forest و شبکه عصبی گراف مکانی-زمانی (ST-GNN) برای پیش‌بینی تعداد مراجعات ماهانه مشتریان ارزیابی گردید. همچنین به منظور تفسیر نتایج و شناسایی عوامل مؤثر بر تقاضا، از روش SHAP و تحلیل اهمیت متغیرها بهره گرفته شد.

تاریخ دریافت: ۱۶ آذر ۱۴۰۴
تاریخ دوری: ۲۹ دی ۱۴۰۴
تاریخ اصلاح: ۱۱ اردیبهشت ۱۴۰۵
تاریخ پذیرش: ۲۵ خرداد ۱۴۰۵

واژگان کلیدی:

داده‌کاوی مکانی
یادگیری ماشین
بانکداری هوشمند
پیش‌بینی تقاضا
تبیین‌پذیری مدل
شبکه عصبی گراف زمانی-مکانی

* نویسنده مسئول

✉ jrsgr@sru.ac.ir

① ۰۵۱-۳۸۶۸۳۹۰۰

یافته‌ها: نتایج تحلیل فضایی نشان داد که تقاضای خدمات بانکی دارای ساختار فضایی معنادار بوده و مقدار شاخص موران‌ای برابر با ۰.۷۱ وجود خودهمبستگی مکانی قوی و تمرکز فضایی تقاضا را تأیید می‌کند. تحلیل خوشه‌های مکانی نیز تمرکز تقاضا را در برخی شهرستان‌های استان نشان داد. نتایج مدل‌سازی بیانگر آن بود که مدل‌های مبتنی بر درخت بهترین عملکرد را در پیش‌بینی تعداد مراجعات ماهانه مشتریان ارائه می‌کنند؛ به‌گونه‌ای که مدل Random Forest با ضریب تعیین (R^2) برابر با ۰.۹۵ و مدل XGBoost با مقدار ۰.۹۴ دقیق‌ترین نتایج را تولید کردند. در مقابل، مدل ST-GNN با وجود توانایی مدل‌سازی روابط مکانی، عملکرد ضعیف‌تری ($R^2 \approx 0.51$) نشان داد که می‌تواند ناشی از محدودیت طول سری زمانی و کمبود داده‌های زمانی پیوسته باشد. نتایج تحلیل‌های تبیین‌پذیر نیز نشان داد که ویژگی‌های مکانی، اقتصادی-اجتماعی و ترانکشی نقش تعیین‌کننده‌ای در شکل‌گیری الگوهای تقاضای خدمات بانکی دارند.

نتیجه‌گیری: یافته‌های پژوهش نشان می‌دهد که ادغام تحلیل مکانی، یادگیری ماشین و روش‌های تبیین‌پذیر می‌تواند چارچوبی مؤثر برای توسعه سامانه‌های تصمیم‌یار در بانکداری هوشمند فراهم کند. چارچوب پیشنهادی علاوه بر بهبود دقت پیش‌بینی تقاضای خدمات بانکی، امکان شناسایی الگوهای فضایی تقاضا و پشتیبانی از تصمیمات مرتبط با توسعه خدمات دیجیتال، مکان‌یابی شعب و تخصیص بهینه منابع را فراهم می‌سازد. محدودیت اصلی پژوهش، کوتاه بودن دوره زمانی داده‌ها و نبود سری‌های زمانی بلندمدت بود که امکان ارزیابی کامل مدل‌های زمانی را محدود کرد. پیشنهاد می‌شود در پژوهش‌های آینده از داده‌های زمانی بلندمدت، داده‌های بلادرنگ و مدل‌های پیشرفته مکانی-زمانی برای توسعه سامانه‌های تصمیم‌یار بانکداری هوشمند استفاده شود.

مقدمه

صنعت بانکداری در دهه‌های اخیر و به‌ویژه در قرن بیست‌ویکم، تحت تأثیر هم‌زمان مجموعه‌ای از روندهای ساختاری از جمله تحول دیجیتال، پیشرفت‌های شتابان فناوری اطلاعات، گسترش کانال‌های بانکداری الکترونیکی و تغییرات بنیادین در انتظارات و رفتار مشتریان، دستخوش دگرگونی عمیقی شده است. این تحولات موجب شده‌اند بانک‌ها به‌تدریج از الگوهای سنتی تصمیم‌گیری مبتنی بر تجربه فردی و قضاوت‌های کارشناسی فاصله گرفته و به سمت رویکردهای داده‌محور و هوشمند حرکت کنند؛ رویکردهایی که در آن‌ها داده به‌عنوان یک دارایی راهبردی و منبع اصلی خلق ارزش تلقی می‌شود [۱،۲]. در این چارچوب، مفهوم بانکداری داده‌محور (Data-Driven Banking) به‌عنوان یک پارادایم غالب مطرح شده است که بر تحلیل داده‌های عظیم، کشف الگوهای پنهان و پشتیبانی نظام‌مند از تصمیم‌گیری‌های عملیاتی و راهبردی تأکید دارد [۳].

مطالعات اخیر نشان می‌دهند که تحول دیجیتال در خدمات مالی به‌طور فزاینده‌ای با بهره‌گیری از تحلیل‌های پیش‌بینانه، هوش مصنوعی و سامانه‌های تصمیم‌یار مبتنی بر داده همراه شده است و این فناوری‌ها به یکی از ارکان اصلی بانکداری هوشمند تبدیل شده‌اند [۴،۵]. بانک‌های پیشرو جهانی، نظیر "JPMorgan Chase" و "HSBC" با به‌کارگیری فناوری‌های پیشرفته هوش مصنوعی و یادگیری ماشین (Machine Learning)، توانسته‌اند از ظرفیت‌های داده‌محوری برای ارائه خدمات شخصی‌سازی‌شده و پاسخ‌گو به نیازهای متنوع مشتریان بهره ببرند. این رویکردها نه تنها به بهبود تجربه و رضایت مشتریان منجر شده‌اند، بلکه از طریق بهینه‌سازی فرآیندهای عملیاتی، کاهش هزینه‌ها و افزایش بهره‌وری، سودآوری و انعطاف‌پذیری نظام‌های بانکی را نیز ارتقا داده‌اند [۳]. مطالعات اخیر نشان می‌دهند که همگرایی هوش مصنوعی، تحلیل

داده‌های عظیم و سیستم‌های تصمیم‌یار هوشمند، در حال تبدیل شدن به یکی از پیشران‌های اصلی تحول دیجیتال در صنعت خدمات مالی است و بانک‌ها به‌طور فزاینده‌ای از مدل‌های یادگیری ماشین برای پیش‌بینی رفتار مشتریان، مدیریت ریسک و بهینه‌سازی خدمات استفاده می‌کنند [۶،۷]. مجموعه این شواهد نشان می‌دهد که بانکداری داده‌محور، زمینه‌ساز گذار از نظام‌های بانکی ایستا به سامانه‌های هوشمند و پویا است که قادرند به‌صورت مستمر و تطبیقی به تغییرات محیطی و انتظارات در حال تحول مشتریان پاسخ دهند. ظهور و بلوغ فناوری‌های هوش مصنوعی و یادگیری ماشین، نقش تعیین‌کننده‌ای در این گذار ایفا کرده‌اند. الگوریتم‌های پیشرفته یادگیری ماشین، به‌ویژه مدل‌های یادگیری عمیق و تقویتی، امکان تحلیل حجم انبوهی از داده‌های ترانکشی، رفتاری و زمینه‌ای را فراهم کرده و بانک‌ها را قادر ساخته‌اند تا رفتار مشتریان را با دقت بالا مدل‌سازی و پیش‌بینی کنند [۸،۹]. مطالعات متعددی نشان داده‌اند که استفاده از این الگوریتم‌ها می‌تواند به بهبود شخصی‌سازی خدمات، افزایش رضایت مشتریان، کاهش ریسک اعتباری و ارتقای بهره‌وری عملیاتی منجر شود [۱۰،۱۱]. در نتیجه، بانک‌های پیشرو جهانی با بهره‌گیری از تحلیل‌های پیش‌بینانه و هوش مصنوعی، توانسته‌اند مزیت رقابتی پایدار در بازارهای مالی پیچیده و پویا ایجاد کنند [۳]. در کنار بُعد داده‌های ترانکشی و رفتاری، بُعد مکانی داده‌ها نیز به‌عنوان یکی از مؤلفه‌های کلیدی در تحلیل سیستم‌های بانکی مطرح شده است. رفتار مشتریان بانکی به‌طور ذاتی وابسته به فضا و مکان است؛ عواملی نظیر توزیع جمعیت، ساختار اقتصادی مناطق، دسترسی مکانی به شعب و رقابت مکانی میان بانک‌ها، تأثیر مستقیمی بر الگوی تقاضای خدمات بانکی دارند [۱۲،۱۳]. داده‌کاوی مکانی و تحلیل‌های مکانی با استفاده از ابزارهای GIS و شاخص‌های خودهمبستگی مکانی، امکان شناسایی خوشه‌های تقاضا و الگوهای مکانی رفتار مشتریان را فراهم

اصلی پژوهش بر تحلیل مکانی الگوهای تقاضا و بهره‌گیری از روش‌های داده‌کاوی مکانی، تحلیل مکانی و مدل‌های یادگیری ماشین مبتنی بر ویژگی‌های مکانی قرار گرفته است. این رویکرد امکان شناسایی ساختارهای مکانی تقاضا، خوشه‌های مکانی و روابط مکانی مؤثر بر رفتار مشتریان را فراهم کرده و مبنای مناسبی برای تصمیم‌سازی در زمینه توسعه و بهینه‌سازی شبکه شعب بانکی ایجاد می‌کند. این شکاف پژوهشی و کاربردی، ضرورت توسعه چارچوب‌هایی یکپارچه و هوشمند را برجسته می‌سازد که بتوانند داده‌های تراکنشی، مکانی و اقتصادی-اجتماعی را به‌صورت هم‌زمان تحلیل کرده و از آن‌ها در پشتیبانی از تصمیم‌گیری‌های راهبردی بانک‌ها استفاده نمایند. پژوهش حاضر در پاسخ به این نیاز، با تمرکز بر ادغام داده‌کاوی مکانی و الگوریتم‌های پیشرفته یادگیری ماشین، به دنبال ارائه مدلی پویا برای تحلیل رفتار مشتریان، پیش‌بینی تقاضای خدمات بانکی و بهینه‌سازی مکان‌یابی شعب است. این رویکرد نه تنها به ارتقای کارایی عملیاتی و بهبود تجربه مشتری کمک می‌کند، بلکه زمینه‌ساز تحقق بانکداری هوشمند بومی‌سازی‌شده و سازگار با شرایط محلی در ایران خواهد بود. برخلاف بسیاری از مطالعاتی که صرفاً بر مقایسه عملکرد الگوریتم‌های یادگیری ماشین متمرکز بوده‌اند، هدف اصلی پژوهش حاضر حل یک مسئله کاربردی در نظام بانکی کشور است. مسئله مورد بررسی، شناسایی الگوهای مکانی تقاضای خدمات بانکی و پشتیبانی از تصمیم‌گیری برای توسعه شبکه شعب و خدمات بانکی در استان مرکزی است. در این راستا، الگوریتم‌های یادگیری ماشین نه به‌عنوان هدف پژوهش، بلکه به‌عنوان ابزارهایی برای استخراج دانش از داده‌ها و پشتیبانی از تصمیمات مکانی مورد استفاده قرار گرفته‌اند.

مبانی نظری و پیشینه پژوهش

یادگیری ماشین در بانکداری

یادگیری ماشین به‌عنوان یکی از ارکان اصلی هوش مصنوعی، نقش فزاینده‌ای در تحول نظام‌های مالی و بانکی ایفا می‌کند. این حوزه با اتکا بر الگوریتم‌هایی که قادر به یادگیری از داده‌ها و بهبود عملکرد خود بدون برنامه‌نویسی صریح هستند، امکان تحلیل حجم عظیمی از داده‌های پیچیده مالی را فراهم ساخته است [۱]. در بانکداری مدرن، یادگیری ماشین به‌طور گسترده در حوزه‌هایی نظیر پیش‌بینی تقاضای خدمات، تحلیل رفتار مشتریان، تشخیص تقلب، ارزیابی ریسک اعتباری و شخصی‌سازی خدمات به کار گرفته می‌شود [۱۱، ۳]. در سال‌های اخیر، الگوریتم‌های یادگیری ماشین به‌ویژه مدل‌های مبتنی بر درخت و روش‌های تقویتی مانند XGBoost، به دلیل توانایی بالا در استخراج الگوهای پیچیده از داده‌های غیرخطی، در حوزه‌های مختلف از جمله تحلیل رفتار مشتریان و پیش‌بینی تقاضا کاربرد گسترده‌ای یافته‌اند [۱۸]. مطالعات تجربی نشان داده‌اند که الگوریتم‌های یادگیری نظارتی مانند رگرسیون جنگل تصادفی (Random Forest Regression) و XGBoost (eXtreme Gradient Boosting) به دلیل توانایی بالا در

می‌کنند و نقش مهمی در تصمیماتی نظیر مکان‌یابی و تخصیص بهینه شعب ایفا می‌نمایند [۱۴، ۱۵].

پژوهش‌های اخیر نشان می‌دهند که ترکیب تحلیل مکانی با الگوریتم‌های یادگیری ماشین، نسبت به رویکردهای تک‌بعدی، دقت بالاتری در پیش‌بینی تقاضا و تصمیم‌سازی مکانی ارائه می‌دهد [۱۳، ۱۶]. به‌ویژه، توسعه مدل‌های زمانی-مکانی پیشرفته مانند شبکه‌های عصبی گراف زمانی-مکانی (ST-GNN) امکان مدل‌سازی هم‌زمان وابستگی‌های مکانی و پویایی‌های زمانی را فراهم کرده و افق‌های جدیدی در تحلیل سیستم‌های پیچیده مالی گشوده است [۱۰].

در ایران، توسعه سریع و گسترده سامانه‌های بانکداری الکترونیکی، اینترنتی و موبایلی طی سال‌های اخیر منجر به تولید حجم عظیمی از داده‌های تراکنشی و رفتاری شده است. بر اساس گزارش‌های رسمی بانک مرکزی جمهوری اسلامی ایران، بیش از ۸۰ درصد تراکنش‌های بانکی کشور از طریق کانال‌های دیجیتال انجام می‌شود که این امر ظرفیت بالقوه قابل‌توجهی برای تحلیل داده‌محور و ارتقای کیفیت خدمات بانکی فراهم می‌آورد (بانک مرکزی جمهوری اسلامی ایران، ۱۴۰۲). با وجود این ظرفیت چشمگیر، بهره‌برداری نظام‌مند از این داده‌ها در حوزه‌هایی نظیر پیش‌بینی تقاضای خدمات بانکی، طراحی سبد خدمات متناسب با ویژگی‌های مکانی-اقتصادی مناطق و مکان‌یابی بهینه شبکه شعب، همچنان محدود و در مراحل ابتدایی قرار دارد. در عمل، بسیاری از تصمیمات مکانی و خدماتی در نظام بانکی ایران همچنان متکی بر روش‌های سنتی، قواعد ایستا و قضاوت‌های ذهنی مدیران است؛ رویکردهایی که توانایی لازم برای شناسایی الگوهای پیچیده و پویای مکانی-زمانی رفتار مشتریان و واکنش مؤثر به تغییرات محیطی را ندارند [۱۷، ۱۵]. این محدودیت‌ها سبب می‌شود که فرصت‌های نهفته در داده‌های عظیم بانکی به‌طور کامل مورد بهره‌برداری قرار نگیرد و در نتیجه، ناکارآمدی در تخصیص منابع، افزایش هزینه‌های عملیاتی و کاهش سطح رضایت مشتریان رخ دهد. مطالعات اخیر نیز تأکید می‌کنند که گذار از رویکردهای سنتی به چارچوب‌های داده‌محور و مبتنی بر یادگیری ماشین، پیش‌شرط اساسی برای بهبود تصمیم‌گیری‌های راهبردی در بانکداری نوین است [۸]. با توجه به ماهیت جغرافیایی خدمات بانکی، توزیع تقاضا برای خدمات مالی در فضا یکنواخت نبوده و تحت تأثیر عواملی نظیر تراکم جمعیت، ساختار اقتصادی مناطق، سطح دسترسی به خدمات، الگوی استقرار شعب و ویژگی‌های اجتماعی-اقتصادی قرار می‌گیرد. از این‌رو، تحلیل بسیاری از مسائل مرتبط با برنامه‌ریزی و مدیریت شبکه شعب بانکی مستلزم توجه هم‌زمان به ابعاد مکانی و داده‌های رفتاری مشتریان است. هرچند داده‌های مورد استفاده در این پژوهش بازه زمانی ۱۴۰۰ تا ۱۴۰۲ را پوشش می‌دهند، اما محدودیت طول سری زمانی و نبود داده‌های پیوسته زمانی در سطح شعب و مشتریان، امکان توسعه و ارزیابی مدل‌های زمانی کامل را محدود ساخته است. بنابراین، تمرکز

می‌کنند که ادغام تحلیل مکانی با مدل‌های پیش‌بینانه، نسبت به رویکردهای سنتی مکان‌یابی، نتایج کارآمدتری در صنایع خدماتی از جمله بانکداری به همراه دارد [۱۵]. با این حال، بخش عمده‌ای از این مطالعات در بستر کشورهای توسعه‌یافته انجام شده و بررسی‌های بومی‌سازی‌شده، به‌ویژه در اقتصادهای در حال توسعه، همچنان محدود است. در این راستا، ظهور رویکردهای نوین مانند GeoAI که ترکیبی از هوش مصنوعی و تحلیل‌های مکانی است، در قلب چارچوب‌های یکپارچه برای کشف الگوهای مکانی پیچیده به کار گرفته شوند و قابلیت پیش‌بینی و تصمیم‌سازی مکانی به شکل معناداری افزایش یابد [۲۰، ۲۲].

شبکه‌های عصبی گراف زمانی مکانی

در سال‌های اخیر، شبکه‌های عصبی گراف (Graph neural networks) به‌عنوان رویکردی نوین برای مدل‌سازی داده‌های ساختاریافته و وابسته معرفی شده‌اند. در این چارچوب، واحدهای مکانی (مانند مناطق یا شعب بانکی) به‌صورت گره‌ها و روابط مکانی یا عملکردی میان آن‌ها به‌صورت یال‌ها نمایش داده می‌شوند. توسعه مدل‌های ST-GNN امکان تحلیل هم‌زمان وابستگی‌های مکانی و پویایی‌های زمانی را فراهم کرده است [۱۰]. در سال‌های اخیر، معماری‌های پیشرفته GNN و Graph Representation Learning به یکی از مهم‌ترین حوزه‌های تحقیقاتی در تحلیل سیستم‌های پیچیده تبدیل شده‌اند و کاربرد گسترده‌ای در حمل‌ونقل، انرژی، خدمات شهری و تحلیل تقاضا یافته‌اند [۲۳، ۲۴]. از اینرو، پژوهش‌ها نشان می‌دهند که مدل‌های ST-GNN در صورت دسترسی به داده‌های زمانی پیوسته و ساختار شبکه‌ای غنی، می‌توانند در پیش‌بینی پدیده‌هایی که به‌طور هم‌زمان به فضا و زمان وابسته‌اند، مانند تقاضای خدمات، ترافیک شهری و مصرف انرژی، عملکرد قابل توجهی ارائه دهند [۲۵]. در حوزه بانکداری، استفاده از این مدل‌ها هنوز در مراحل اولیه قرار دارد، اما نتایج اولیه حاکی از آن است که ST-GNN می‌تواند روابط پیچیده میان مناطق مختلف، رقابت مکانی شعب و تغییرات زمانی رفتار مشتریان را به‌صورت یکپارچه مدل‌سازی کند [۳]. این قابلیت، ST-GNN را به گزینه‌ای مناسب برای پیش‌بینی تقاضای خدمات بانکی در مقیاس مکانی-زمانی تبدیل می‌کند.

مدل مفهومی و فرضیات پژوهش

مدل مفهومی پژوهش حاضر مطابق شکل (۱) بر پایه یک چارچوب یکپارچه داده‌محور طراحی شده است که داده‌های تراکنشی، مکانی و اقتصادی-اجتماعی مشتریان و شعب بانک را در بازه زمانی ۱۴۰۰ تا ۱۴۰۲ به‌عنوان ورودی دریافت می‌کند. پس از مرحله جمع‌آوری، یکپارچه‌سازی و پیش‌پردازش داده‌ها، اطلاعات وارد لایه تحلیل مکانی می‌شوند. در این مرحله، الگوهای مکانی تقاضای خدمات بانکی و ساختارهای مکانی آن با استفاده از روش‌هایی نظیر Moran's، DBSCAN و LISA شناسایی و تحلیل می‌گردند.

مدل‌سازی روابط غیرخطی و مقاومت در برابر بیش‌برازش، عملکرد قابل توجهی در پیش‌بینی رفتار مالی مشتریان دارند [۸]. از سوی دیگر، شبکه‌های عصبی عمیق، به‌ویژه معماری‌های بازگشتی نظیر LSTM، در مدل‌سازی وابستگی‌های زمانی داده‌های تراکنشی عملکرد برتری از خود نشان داده‌اند [۹]. پژوهش Yin و Zhou (2021) نشان می‌دهد که استفاده از مدل‌های عمیق می‌تواند دقت پیش‌بینی الگوهای مصرف مالی را به‌طور معناداری افزایش دهد. با وجود این، بخش قابل توجهی از پژوهش‌های موجود در بانکداری داده‌محور، تمرکز خود را صرفاً بر داده‌های تراکنشی یا رفتاری قرار داده‌اند و کمتر به بُعد مکانی داده‌ها توجه داشته‌اند. این در حالی است که رفتار مشتریان بانکی به‌شدت متأثر از زمینه‌های مکانی و ساختار مکانی محیط‌های اقتصادی-اجتماعی است [۱۰]. در سال‌های اخیر، مطالعات متعددی نشان داده‌اند که انتخاب مدل یادگیری ماشین باید متناسب با ساختار داده، حجم نمونه و غنای زمانی-مکانی انجام شود و استفاده از مدل‌های پیچیده‌تر لزوماً به بهبود عملکرد منجر نمی‌شود [۱۹]. از این‌رو، ترکیب یادگیری ماشین با تحلیل‌های مکانی، به‌عنوان گامی ضروری برای ارتقای دقت و کارایی مدل‌های پیش‌بینانه بانکی مطرح می‌شود.

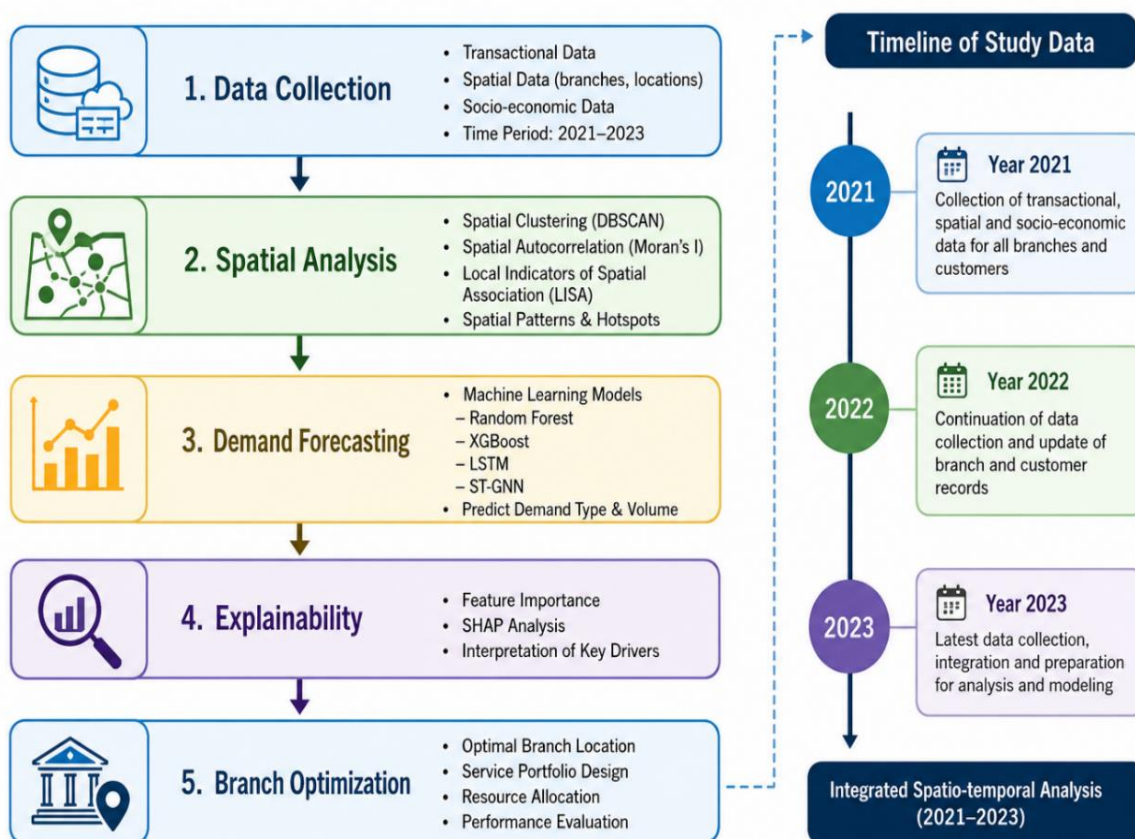
داده‌کاوی مکانی و تحلیل مکانی

داده‌کاوی مکانی شاخه‌ای تخصصی از داده‌کاوی است که به استخراج الگوها، روابط و ساختارهای پنهان از داده‌های دارای مؤلفه مکانی می‌پردازد. برخلاف داده‌های سنتی، داده‌های مکانی دارای وابستگی مکانی هستند؛ بدین معنا که مقادیر مشاهده‌شده در یک مکان، به مقادیر مکان‌های مجاور وابسته‌اند [۱۲]. مطالعات جدید نشان می‌دهد که مدل‌های مکانی یادگیری ماشین (Learning Spatial Machine) قادرند وابستگی‌های مکانی و ناهمگنی مکانی را بهتر از مدل‌های آماری کلاسیک مدل‌سازی کنند و در مسائل تخصیص خدمات و برنامه‌ریزی منطقه‌ای عملکرد بهتری ارائه دهند [۲۰، ۲۱]. این ویژگی موجب می‌شود که روش‌های کلاسیک داده‌کاوی، بدون در نظر گرفتن ساختار مکانی، نتوانند الگوهای واقعی را به‌درستی شناسایی کنند. در این زمینه، شاخص‌هایی نظیر Moran's I و Local Indicators of Spatial Association) به‌طور گسترده برای سنجش خودهمبستگی مکانی و شناسایی خوشه‌های مکانی به کار رفته‌اند [۱۲]. پژوهش‌های متعددی نشان داده‌اند که استفاده از این شاخص‌ها در تحلیل تقاضای خدمات، امکان شناسایی مناطق همگن از نظر رفتار مصرفی را فراهم می‌کند [۱۴]. در حوزه بانکداری، داده‌کاوی مکانی به‌ویژه در مسائل مکان‌یابی و تخصیص شعب بانکی اهمیت یافته است؛ چراکه توزیع نامناسب شعب می‌تواند منجر به افزایش هزینه‌ها و کاهش سطح دسترسی مشتریان شود [۱۷]. مطالعه Li و همکاران در ۲۰۲۰، نشان داد که ترکیب سیستم‌های اطلاعات جغرافیایی (GIS) با الگوریتم‌های یادگیری ماشین می‌تواند مکان‌های بهینه شعب بانکی را با دقت بالا شناسایی کرده و پوشش خدمات را به‌طور معناداری افزایش دهد [۱۳]. همچنین، Rahman و Shukla در تحقیق دیگری در ۲۰۲۲ تأکید

مکانی، الگوریتم‌های پیشرفته یادگیری ماشین و روش‌های تبیین‌پذیری می‌تواند دقت پیش‌بینی تقاضای خدمات بانکی را به‌طور معناداری افزایش داده و کیفیت تصمیمات مرتبط با مکان‌یابی و توسعه شبکه شعب را بهبود بخشد. همچنین فرض می‌شود که ویژگی‌های مکانی و اقتصادی-اجتماعی مناطق بر حجم و نوع تقاضای خدمات بانکی اثرگذار هستند، داده‌های تراکنشی تاریخی ظرفیت مناسبی برای پیش‌بینی رفتار آتی مشتریان دارند، عملکرد مدل‌های پیش‌بینی تحت تأثیر ساختار مکانی و زمانی داده‌ها قرار می‌گیرد و استفاده از نتایج پیش‌بینی در فرآیند تصمیم‌گیری مکانی می‌تواند به افزایش پوشش خدمات، کاهش هزینه‌های عملیاتی و ارتقای کارایی شبکه بانکی منجر شود. در مجموع، مدل مفهومی ارائه‌شده چارچوبی منسجم برای اتصال داده‌ها، تحلیل‌های مکانی، پیش‌بینی تقاضا، تبیین‌پذیری مدل‌ها و تصمیم‌سازی مکانی فراهم می‌آورد و می‌تواند مبنایی برای توسعه سامانه‌های تصمیم‌یار هوشمند در بانکداری داده‌محور و بانکداری هوشمند باشد.

در ادامه، خروجی‌های تحلیل مکانی به همراه ویژگی‌های تراکنشی و اقتصادی-اجتماعی در لایه پیش‌بینی تقاضا مورد استفاده قرار می‌گیرند. برای این منظور، مدل‌های یادگیری ماشین و یادگیری عمیق شامل Random Forest، XGBoost، LSTM و ST-GNN به کار گرفته می‌شوند تا حجم تقاضا، نوع خدمات و کنال‌های بانکی ترجیحی مشتریان پیش‌بینی شود. سپس، نتایج حاصل با بهره‌گیری از روش‌های تبیین‌پذیری مانند تحلیل اهمیت ویژگی‌ها و SHAP تفسیر شده و عوامل مؤثر بر رفتار مشتریان شناسایی می‌شوند.

در نهایت، خروجی‌های پیش‌بینی و تبیین‌پذیری به‌عنوان ورودی لایه تصمیم‌سازی و بهینه‌سازی شبکه شعب مورد استفاده قرار می‌گیرند تا از آن‌ها در مکان‌یابی شعب، تخصیص خدمات و برنامه‌ریزی توسعه شبکه بانکی بهره گرفته شود. همچنین، به‌منظور نمایش بعد زمانی پژوهش، داده‌های مورد استفاده در قالب یک خط زمانی شامل سال‌های ۲۰۲۱، ۲۰۲۲ و ۲۰۲۳ در مدل مفهومی نمایش داده شده‌اند که بیانگر استفاده از اطلاعات گردآوری‌شده در طول دوره مطالعه است. بر این اساس، فرضیه اصلی پژوهش بیان می‌کند که ادغام تحلیل‌های



شکل ۱: مدل مفهومی پژوهش

Fig. 1: Conceptual Modeling of research

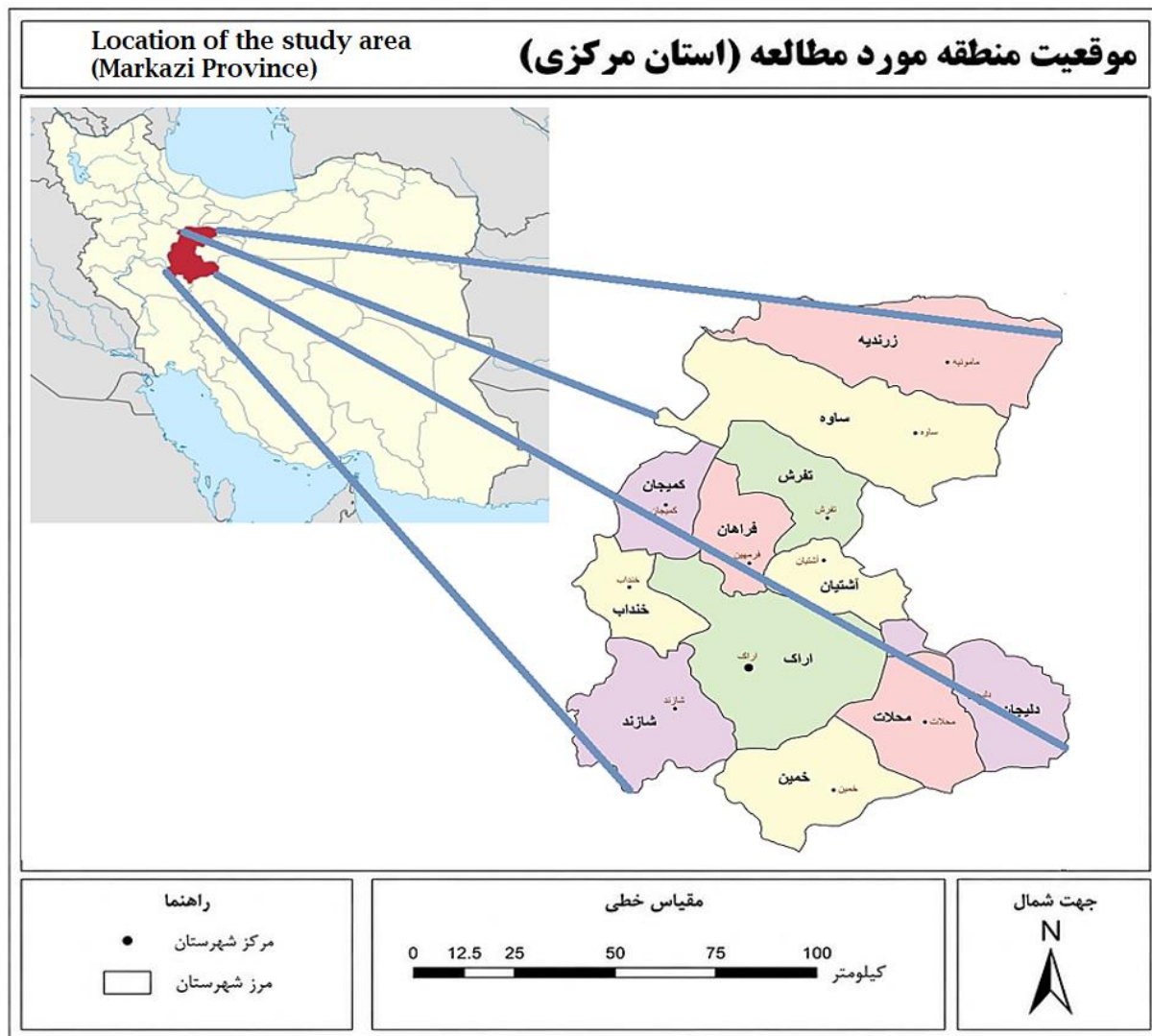
منطقه مورد مطالعه

۱۴۰۲، این استان را به بستری مناسب برای تحلیل مکانی تقاضای خدمات بانکی و ارزیابی مدل‌های پیش‌بینی تبدیل کرده است.

روش‌شناسی پژوهش

پژوهش حاضر از نظر هدف، کاربردی و از حیث ماهیت روش‌شناختی، تحلیلی-داده‌محور است و در چارچوب رویکردهای نوین تصمیم‌سازی مبتنی بر داده و هوش مصنوعی انجام شده است. هدف اصلی این پژوهش، توسعه و ارزیابی یک چارچوب یکپارچه برای تحلیل رفتار مشتریان، پیش‌بینی تقاضای خدمات بانکی و بهینه‌سازی مکان‌یابی شعب، با اتکا بر ادغام داده‌های تراکنشی، مکانی و اقتصادی-اجتماعی است. از این‌رو، روش تحقیق بر پایه مدل‌سازی کمی، یادگیری ماشین و تحلیل‌های مکانی-زمانی طراحی شده است. فلوچارت روش‌شناسی پژوهش مطابق با تصویر شماره ۳ است.

استان مرکزی در مرکز ایران و در محدوده جغرافیایی ۳۳ تا ۳۵ درجه عرض شمالی و ۴۹ تا ۵۱/۵ درجه طول شرقی واقع شده است. این استان با مساحتی حدود ۲۹ هزار کیلومتر مربع و مرکزیت شهر اراک، یکی از قطب‌های مهم صنعتی و خدماتی کشور محسوب می‌شود. شکل (۲) موقعیت جغرافیایی استان مرکزی و تقسیمات شهرستانی آن را نشان می‌دهد. جمعیت استان بیش از ۴/۱ میلیون نفر بوده و عمدتاً در شهرستان‌های اراک، ساوه، خمین، محلات و دلیجان متمرکز است. وجود ترکیبی از مناطق شهری، صنعتی و روستایی، الگوهای متنوعی از تقاضای خدمات بانکی را در سطح استان ایجاد کرده است. استان مرکزی به دلیل استقرار صنایع بزرگ، موقعیت ارتباطی مناسب و نقش مؤثر در فعالیت‌های اقتصادی کشور، از تنوع مکانی قابل توجهی در رفتارهای اقتصادی و مالی برخوردار است. همچنین دسترسی به داده‌های واقعی مشتریان و شعب بانک سپه در بازه زمانی ۱۴۰۰ تا



شکل ۲: موقعیت جغرافیایی استان مرکزی در ایران

Fig. 2: Location of the study area (Markazi Province)

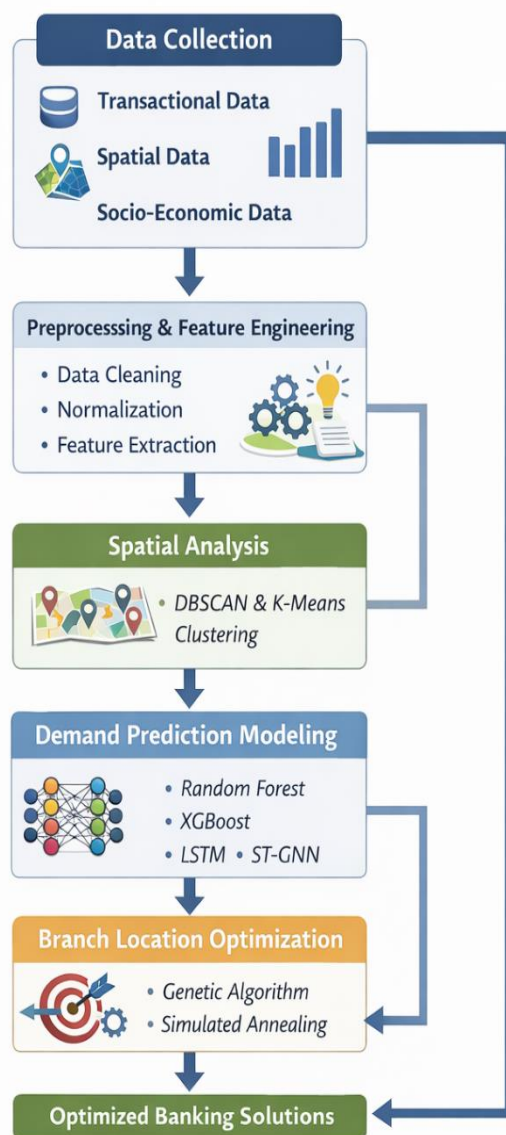
رفتاری مشتری (مانند سن، جنسیت، تحصیلات، شغل، سطح درآمد و تعداد مراجعات ماهانه و همچنین خروجی‌های رفتاری/خدماتی شامل کلنل بانکداری غلب و نوع خدمات دریافتی استخراج گردید. دوم، داده‌های مکانی مبتنی بر GIS شامل مختصات جغرافیایی (Latitude/Longitude) و لایه‌های مرجع مکانی مانند شبکه معابر و الگوهای توزیع جمعیت برای زمینه‌سازی تحلیل مکانی استفاده شد. سوم، داده‌های جمعیتی-اقتصادی (نظیر ساختار فعالیت اقتصادی، سطح اشتغال و شاخص‌های رفاهی) به‌منظور تبیین تفاوت‌های مکانی تقاضا لحاظ شد. این ترکیب چندمنبعی، امکان تحلیل چندبعدی الگوهای تقاضا را فراهم کرده و ریسک سوگیری ناشی از اتکای صرف به یک منبع داده را کاهش می‌دهد.

برای شفاف‌سازی ساختار داده‌های مورد استفاده، نوع متغیرها و منبع استخراج هر یک از آن‌ها، مشخصات متغیرهای اصلی پژوهش در جدول (۱) ارائه شده است. این متغیرها شامل ویژگی‌های رفتاری، خدماتی، اقتصادی و مکانی مشتریان بوده و به‌عنوان ورودی مدل‌های تحلیل مکانی و یادگیری ماشین مورد استفاده قرار گرفته‌اند.

پیش‌پردازش داده‌ها و مهندسی ویژگی‌ها

در این پژوهش، داده‌ها پیش از ورود به مرحله مدل‌سازی، تحت یک خط لوله‌ی پاک‌سازی و استانداردسازی قرار گرفتند تا سازگاری ساختاری، کیفیت آماری و قابلیت بازتولید (Reproducibility) تضمین شود. نخست، ناسازگاری‌های محتمل در نام‌گذاری ستون‌ها از طریق یک لایه‌ی همسان‌سازی شمای داده (Schema Harmonization) رفع شد؛ به‌طوری‌که نگاشت نام‌های جایگزین/اشتباه به نام‌های استاندارد انجام گرفت و در نتیجه، داده‌ی ورودی با ساختار مورد انتظار مدل‌ها یکسان‌سازی شد. در گام بعد، نوع داده‌ها به‌ویژه برای متغیرهای عددی (سن، مختصات جغرافیایی، تعداد مراجعات ماهانه و سایر فیلدهای کمی) به‌صورت نظام‌مند تبدیل و مقادیر غیرقابل تفسیر به مقادیر گم‌شده‌ی استاندارد تبدیل شدند. سپس، رکوردهای فاقد مختصات مکانی معتبر (Latitude/Longitude) حذف شدند تا تحلیل‌های مکانی و گراف‌محور بر پایه‌ی داده‌ی مکانی قابل اتکا انجام گیرد.

پس از پاک‌سازی، پیش‌پردازش ویژگی‌ها با هدف تولید ورودی مناسب برای مدل‌های کلاسیک و عمیق انجام شد. متغیرهای عددی با روش مقیاس‌گذار استاندارد (StandardScaler) استانداردسازی نرمال‌سازی شدند و متغیرهای طبقه‌ای با کدگذاری یک به چند (One-Hot Encoding) تبدیل به نمایش برداری گردیدند تا هم‌زمان امکان استفاده در مدل‌های درختی و شبکه‌های عصبی فراهم شود. برای افزایش ظرفیت تبیینی مدل‌ها، علاوه بر ویژگی‌های خام، یک شاخص ترکیبی به‌عنوان «شاخص قدرت اقتصادی محلی (Local Economic Index proxy)» تعریف شد.



شکل ۳: فلوچارت روش‌شناسی پژوهش؛ مراحل گردآوری داده، پیش‌پردازش،

مهندسی ویژگی، تحلیل‌های مکانی، مدل‌سازی و ارزیابی نتایج

Fig. 3: Flowchart of the research methodology; stages of data collection, preprocessing, feature engineering, spatial analyses, modeling, and results evaluation.

گردآوری داده‌ها و منابع اطلاعاتی

در این پژوهش، از قسمتی از داده‌های شعب بانک سپه استان مرکزی استفاده شد. داده‌های مورد استفاده شامل اطلاعات تراکنشی، جمعیتی، اقتصادی و مکانی از مشتریان و شعب بانک‌های سپه استان مرکزی در بازه زمانی ۱۴۰۰ تا ۱۴۰۲ می‌باشد. داده‌های مورد استفاده در این پژوهش از چند منبع مکمل گردآوری شد تا پوشش جامعی از ابعاد رفتاری، مکانی و زمینه‌ای تقاضای خدمات بانکی فراهم شود. نخست، داده‌های عملیاتی بانک شامل ویژگی‌های جمعیت‌شناختی و

جدول ۱: مشخصات متغیرهای مورد استفاده در پژوهش

Table 1: Characteristics of the variables used in the research

متغیر VARIABLE	شرح متغیر VARIABLE DESCRIPTION	نوع داده DATA TYPE	منبع داده DATA SOURCE
VPM (VISITS PER MONTH) بازدید در ماه	تعداد مراجعات ماهانه مشتری number of monthly customer visits	عددی Numerical	تراکنش‌ها و سوابق بانکی bank transactions and records
DC (DOMINANT CHANEL)	کانال بانکداری غالب (حضوری، اینترنتی، همراه بانک) dominant banking channel (Internet, Person, Mobile banking)	طبقه‌ای (Categorical)	سامانه بانکداری banking system
SERVICE TYPE (ST)	نوع خدمت دریافتی type of service received	طبقه‌ای (Categorical)	سامانه بانکداری banking system
AGE	سن مشتری customer age	عددی Numerical	پروفایل مشتری customer profile
GENDER	جنسیت مشتری customer gender	طبقه‌ای Numerical	پروفایل مشتری customer profile
EDUCATION	سطح تحصیلات education level	طبقه‌ای Numerical	پروفایل مشتری customer profile
OCCUPATION	شغل مشتری customer's job	طبقه‌ای Numerical	پروفایل مشتری customer profile
INCOME	سطح درآمد مشتری customer income level	عددی / ترتیبی	پروفایل مشتری customer profile
LATITUDE	عرض جغرافیایی latitude	مکانی (Spatial)	مختصات شعب و مشتریان / GIS
LONGITUDE	طول جغرافیایی longitude	مکانی (Spatial)	مختصات شعب و مشتریان / GIS
RESIDENCE TYPE	نوع محل سکونت (شهر، روستا، کلان‌شهر) type of residence (city-village-metropolis)	طبقه‌ای Categorical	اطلاعات مکانی مشتری
BRANCH	شعبه بانکی مربوطه relevant bank branch	طبقه‌ای	سامانه بانکداری

پیاده‌سازی نهایی لحاظ شده و مبنای اعتبار استنتاج‌های بعدی در بخش نتایج است. به‌منظور نمایش ساختار داده، نمونه‌ای از رکوردهای مشتریان و شعب در جدول ۲ ارائه می‌شود. شکل ۴ پراکنش مکانی مشتریان و شعب بانکی را در استان مرکزی نشان می‌دهد. تمرکز بالاتر مشتریان در شهرستان‌های اراک، ساوه و خمین بیانگر وجود کانون‌های اصلی تقاضای خدمات بانکی است. این الگو نشان می‌دهد که توزیع تقاضا در سطح استان یکنواخت نبوده و تحلیل‌های مکانی می‌توانند در شناسایی مناطق دارای ظرفیت توسعه خدمات بانکی نقش مؤثری ایفا کنند.

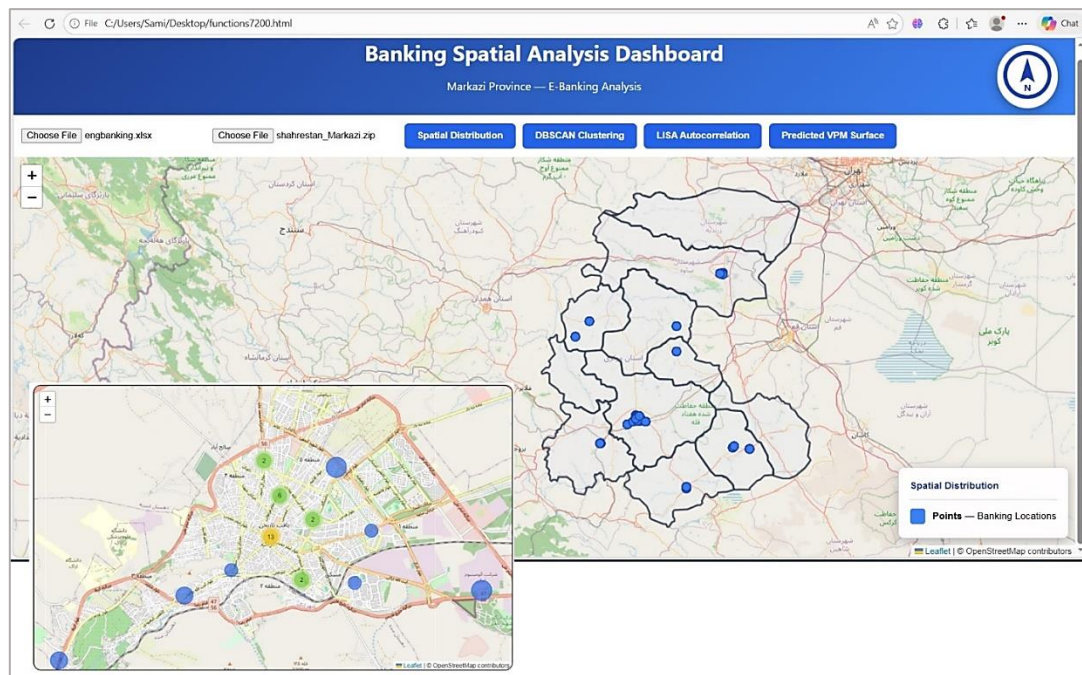
مجموعه داده گردآوری‌شده شامل ۷۰۰۰ رکورد و ۱۰ ویژگی اصلی بود. نسبت داده‌های آموزش و آزمون برابر با ۸۰ به ۲۰ در نظر گرفته شد. برای کنترل اثر عدم توازن احتمالی کلاس‌ها در اهداف طبقه‌بندی DC و ST، علاوه بر Accuracy از معیار F1-macro نیز استفاده شد.

این شاخص بر پایه متغیرهای مستقل اقتصادی-جمعیت‌شناختی و مکانی، مانند سطح درآمد، نوع شغل، محل سکونت و سایر ویژگی‌های زمینه‌ای ساخته شد. با توجه به چندسناریویی بودن مسئله، طراحی و استفاده از این شاخص در هر سناریوی مدل‌سازی به‌صورت هدف‌محور انجام گرفت؛ به این معنا که اگر هر یک از متغیرهای تعداد مراجعه/تعامل ماهانه (VPM)، نوع کانال بانکداری (DC)، نوع خدمت (ST)، به‌عنوان متغیر هدف تعریف می‌شد، آن متغیر از ماتریس ویژگی‌ها و از فرایند ساخت ویژگی‌های ترکیبی حذف می‌گردید. در مقابل، استفاده از این متغیرها در سایر سناریوهایی که نقش متغیر مستقل داشتند مجاز در نظر گرفته شد. این رویکرد نظام‌مند در تفکیک متغیرهای ورودی و هدف، تضمین می‌کند که هیچ‌گونه نشست اطلاعات در فرایند یادگیری رخ نداده و نتایج به‌دست‌آمده از اعتبار آماری و قابلیت تعمیم‌پذیری برخوردار باشند. این اصل به‌عنوان یکی از الزامات کلیدی در طراحی مدل‌های داده‌محور در این پژوهش و در

جدول ۲: نمونه‌ای از داده‌های مشتریان و شعب (ویژگی‌های جمعیت‌شناختی، اقتصادی، رفتاری و مکانی)

Table 2: Sample customer and branch data (demographic, economic, behavioral, and location characteristics)

شناسه مشتری	سن	جنسیت	تحصیلات	شغل	درآمد (میلیون)	دفعات مراجعه ماهانه	کانال بانکداری غالب	نوع خدمات دریافتی	محل سکونت	شعبه بانک
مشتري	Age	Gender	Education	job	(Income)	Monthly visits	Dominant banking channel	Type of service received	Location	Bank branch
1	18	مرد	دیپلم	دامدار	۲۵ تا ۱۵	6	حضور	افتتاح حساب و خدمات کارت	روستا	مهاجران Mohajeran
2	19	مرد	فوق دیپلم	کارگر	۲۵ تا ۱۵	8	تلفن بانک	تسهیلات وام	شهر	ساوه Saveh
3	24	زن	فوق دیپلم	کارگر	۲۵ تا ۱۵	3	حضور	تسهیلات وام	شهر	سنجان Senejan
4	27	زن	لیسانس	کارمند	۲۵ تا ۱۵	7	همراه بانک	سپرده گذاری	کلانشهر	اراک میدان شهدا Square Shohada
5	34	مرد	لیسانس	کشاورز	۵۰ تا ۲۵	3	حضور	افتتاح حساب و خدمات کارت	کلانشهر	ولیعصر اراک Valiasre
6	36	زن	لیسانس	کارمند	۵۰ تا ۲۵	6	حضور	سپرده گذاری	شهر	مطهری ساوه Motahary
7	40	زن	لیسانس	وکیل	بالای ۵۰	5	همراه بانک	تسهیلات وام	کلانشهر	شهید رجایی اراک Rajae
8	45	مرد	فوق لیسانس	صاحب کسب و کار	بالای ۵۰	20	اینترنتی	خدمات چک و چکاوک	کلانشهر	باهر اراک Bahunar
9	62	مرد	فوق لیسانس	کارمند	۵۰ تا ۲۵	5	تلفن بانک	سپرده گذاری	کلانشهر	بازار اراک bazar
10	66	مرد	دکتر	مدیر	بالای ۵۰	15	همراه بانک	ضمانت نامه و ال سی	کلانشهر	بهشتی اراک Beheshty



شکل ۴: توزیع مکانی شعب بانکهای سپه در منطقه مورد مطالعه

Fig. 4: Spatial distribution of Sepah Bank branches in the study area

مدل سازی پیش‌بینی تقاضای خدمات بانکی

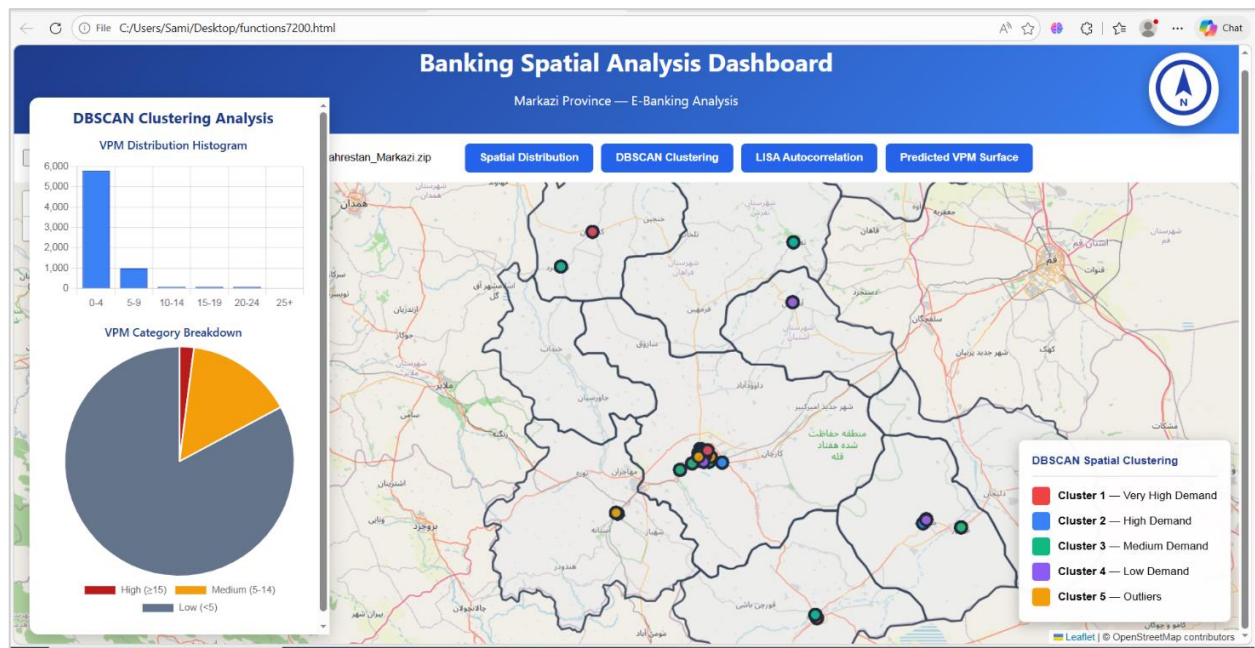
به منظور مدل‌سازی تقاضا و رفتار خدماتی مشتریان، سه سناریوی هدف تعریف شد: (۱) به عنوان هدف رگرسیونی برای مدل‌سازی شدت/حجم تعامل، (۲) کانال غالب خدمت DC به عنوان هدف طبقه‌بندی برای پیش‌بینی ترجیح کانال خدمت، و (۳) نوع خدمت بانکداری ST به عنوان هدف طبقه‌بندی برای پیش‌بینی نوع خدمت غالب. برای هر سناریو، یک ماتریس ویژگی سازگار با همان هدف ساخته شد و پیش‌پردازش هدف‌محور اعمال گردید تا هم‌زمان قابلیت مقایسه منصفانه بین مدل‌ها حفظ شود. تقسیم داده‌ها به صورت ۸۰٪ آموزش و ۲۰٪ آزمون انجام شد. همچنین برای جلوگیری از بیش‌برازش و ارزیابی پایدار عملکرد مدل‌ها از اعتبارسنجی متقاطع K-Fold با K=5 استفاده گردید.

توجیه انتخاب مدل‌ها

انتخاب مدل‌های پیش‌بینی در این پژوهش با هدف شناسایی مناسب‌ترین ابزار برای پشتیبانی از تصمیم‌گیری مکانی در بانکداری انجام شده است. بنابراین مقایسه مدل‌ها صرفاً یک ارزیابی فنی محسوب نمی‌شود، بلکه ابزاری برای تعیین میزان قابلیت استفاده از هر مدل در پیش‌بینی تقاضای خدمات بانکی و پشتیبانی از تصمیمات مرتبط با توسعه شبکه شعب و تخصیص خدمات در مناطق مختلف استان مرکزی است. از اینرو، مدل‌های مورد استفاده در این پژوهش با هدف مقایسه رویکردهای کلاسیک یادگیری ماشین و روش‌های نوین یادگیری عمیق مبتنی بر ساختارهای مکانی انتخاب شدند.

تحلیل مکانی و خوشه‌بندی مکانی

برای شناسایی ساختار مکانی تقاضا و تحلیل ناهمگنی مکانی رفتار بانکی، تحلیل مکانی در محیط GIS و با رویکرد خوشه‌بندی مکانی انجام شد. در این مرحله، از دو الگوریتم مکمل استفاده شد تا هم قابلیت شناسایی خوشه‌های نامنظم و نقاط پرت و هم امکان تقسیم‌بندی همگن برای تفسیر مدیریتی فراهم گردد. به طور مشخص، روش خوشه‌بندی مکانی مبتنی بر چگالی کاربردها همراه با نوین (Density-Based Spatial Clustering of Applications with Noise) DBSCAN برای استخراج خوشه‌های با شکل نامنظم و تمایز نقاط پرت به کار رفت؛ این الگوریتم به‌ویژه در داده‌های مکانی ناهمگن، به دلیل اتکاء به چگالی و عدم نیاز به تعیین تعداد خوشه‌ها از پیش، مناسب است. در کنار آن، K-Means برای تفکیک مناطق به خوشه‌های همگن‌تر مبتنی بر ویژگی‌های مکانی-اقتصادی استفاده شد تا یک تصویر «تقسیم‌بندی‌شده و قابل سیاست‌گذاری» از مناطق حاصل شود. خروجی خوشه‌بندی به عنوان یک لایه‌ی تبیینی در تحلیل نتایج پیش‌بینی و نیز در استدلال‌های مکان‌یابی مورد استفاده قرار می‌گیرد؛ به این معنا که تفاوت‌های عملکرد مدل‌ها و الگوهای خطا در سناریوهای مختلف، با زمینه‌ی مکانی/اقتصادی هر خوشه تفسیر می‌شود. نتایج شکل ۵ نشان می‌دهد که مشتریان در چند خوشه مکانی متراکم سازمان یافته‌اند و تعدادی نقاط پراکنده نیز به عنوان نوین شناسایی شده‌اند. این خوشه‌ها می‌توانند مبنای طراحی خدمات منطقه‌ای و برنامه‌ریزی توسعه شبکه شعب قرار گیرند.



در داشبورد تحلیل مکانی بانکداری نوین DBSCAN شکل ۵: خوشه‌بندی مکانی تقاضای بانکی با استفاده از الگوریتم
Fig. 5: Spatial clustering of banking demand using the DBSCAN algorithm in the Modern Banking Spatial Analysis Dashboard

مبتنی بر درخت تصمیم، از تنظیم‌گری و بهینه‌سازی گرادینتی برای افزایش پایداری و دقت پیش‌بینی استفاده می‌کند [۱۸]. مقایسه‌ای اخیر نشان داده‌اند که مدل‌های مبتنی بر گرادینت بوسستینگ همچنان در بسیاری از مسائل داده‌های جدولی (Tabular Data) عملکردی رقابتی و حتی برتر از بسیاری از مدل‌های عمیق ارائه می‌کنند [۲۶].

دوم، مدل‌های عمیق شامل حافظه بلند-کوتاه مدت (Long Short-Term Memory) و ST-GNN که به ترتیب برای نمایش توالی‌محور و نمایش مکانی مبتنی بر گراف به کار گرفته شدند. در پیاده‌سازی LSTM، به دلیل ماهیت داده‌ی مقطعی و فقدان سری زمانی حقیقی، داده به صورت توالی طول یک (SeqLen=1) به مدل ارائه شد؛ بنابراین نقش LSTM در این مطالعه بیشتر به عنوان یک «مدل عمیق مقایسه‌ای» تفسیر می‌شود تا اینکه مدل‌سازی وابستگی زمانی واقعی در نظر گرفته شود. در واقع، با توجه به مقطعی بودن داده‌ها و عدم وجود سری زمانی واقعی، استفاده از LSTM با طول توالی ۱ صرفاً به عنوان یک baseline مقایسه‌ای در نظر گرفته شد و قابلیت مدل‌سازی وابستگی زمانی در این مطالعه محدود است. این موضوع به عنوان یکی از محدودیت‌های روش شناختی پژوهش در نظر گرفته می‌شود. در مقابل، ST-GNN برای وارد کردن «ساختار همجواری مکانی» به مدل طراحی شد.

مدل‌های گرافی مانند GCN برای مدل‌سازی وابستگی‌های ساختاری میان گره‌ها توسعه یافته‌اند و در سال‌های اخیر در مسائل مکانی و شبکه‌ای کاربرد گسترده‌ای یافته‌اند. به ویژه، مدل‌های گرافی زمانی-مکانی (ST-GNN) قادر به ترکیب اطلاعات مکانی و زمانی برای تحلیل پدیده‌های پیچیده هستند، هرچند کارایی آن‌ها به شدت وابسته به

مدل‌های Random Forest و XGBoost به دلیل توانایی بالا در مدل‌سازی روابط غیرخطی، پایداری در برابر بیش‌برازش، عملکرد مناسب در داده‌های جدولی و قابلیت مدیریت هم‌زمان متغیرهای عددی و طبقه‌ای انتخاب شدند. این مدل‌ها در بسیاری از مطالعات پیش‌بینی رفتار مشتریان و تحلیل تقاضا عملکرد قابل اعتمادی نشان داده‌اند. علاوه بر این، مدل LSTM به منظور بررسی امکان بهره‌گیری از الگوهای ترتیبی نهفته در داده‌های رفتاری مشتریان مورد استفاده قرار گرفت. هرچند داده‌های این پژوهش ماهیت کاملاً سری زمانی ندارند، استفاده از LSTM امکان مقایسه عملکرد مدل‌های عمیق با مدل‌های مبتنی بر درخت را فراهم می‌کند. همچنین، مدل ST-GNN با هدف ارزیابی تأثیر وابستگی‌های مکانی و ساختارهای گرافی در پیش‌بینی تقاضای خدمات بانکی به کار گرفته شد. این مدل قادر است روابط همجواری و اثرات سرریز مکانی میان واحدهای جغرافیایی را در فرآیند یادگیری لحاظ کند. مقایسه هم‌زمان این دو خانواده از مدل‌ها امکان ارزیابی این موضوع را فراهم می‌سازد که آیا بهره‌گیری از مدل‌های عمیق آگاه مکانی (Spatially-Aware Deep Learning) می‌تواند نسبت به روش‌های متداول یادگیری ماشین، بهبود معناداری در دقت پیش‌بینی و پشتیبانی از تصمیم‌گیری‌های مکانی ایجاد کند یا خیر.

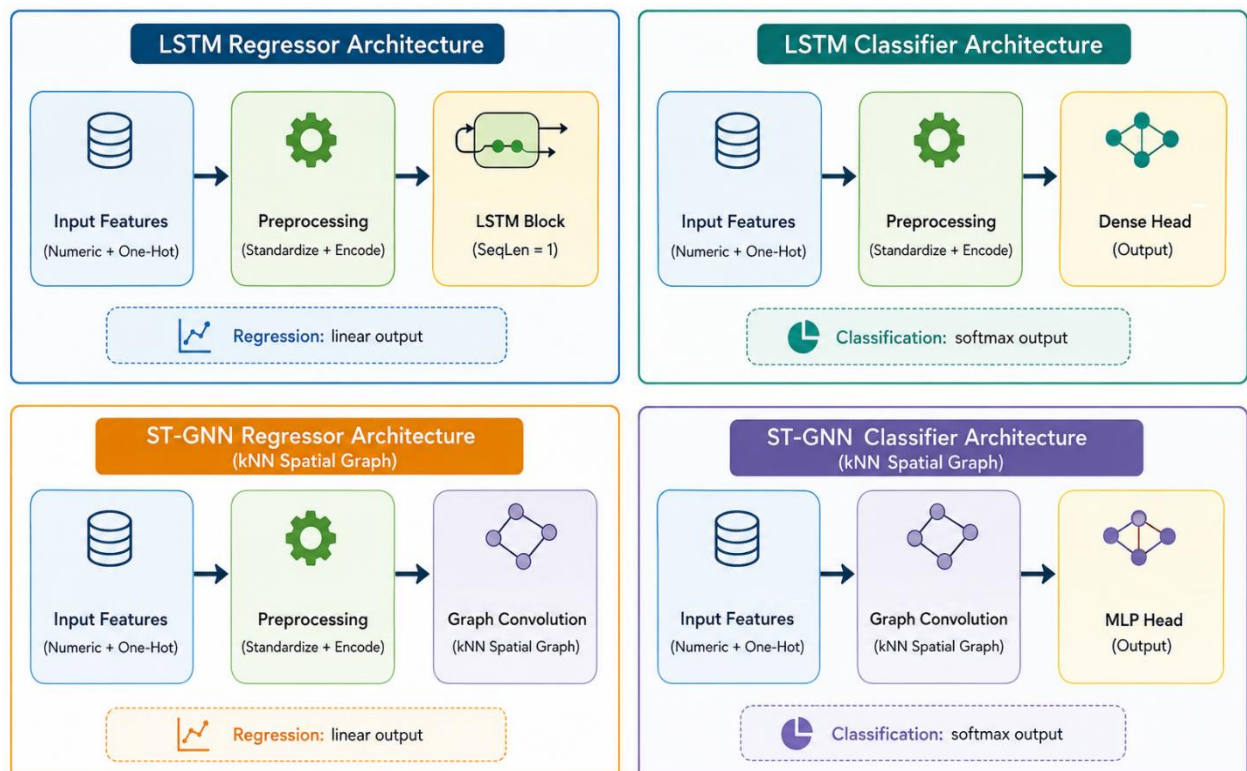
در سطح الگوریتم‌ها، دو خانواده مدل مورد مقایسه قرار گرفت. نخست، مدل‌های کلاسیک مبتنی بر درخت شامل رگرسیون جنگل تصادفی و RFR و XGBoost به کار گرفته شدند. این مدل‌ها به دلیل توانایی بالا در مدل‌سازی روابط غیرخطی، کنترل نسبی بیش‌برازش و کارایی مناسب در داده‌های ترکیبی عددی-طبقه‌ای، در مسائل پیش‌بینی و داده‌کاوی کاربرد گسترده‌ای دارند. به ویژه، XGBoost به عنوان یک روش تقویتی

پیش‌بینی تقاضا با مدل‌های مکان‌یابی خدمات یکی از روندهای نوین در مکان‌یابی هوشمند محسوب می‌شود که امکان تخصیص بهینه منابع و توسعه شبکه خدمات را بر اساس شواهد داده‌محور فراهم می‌کند [۲۹]. در این فرموله‌سازی، تابع هدف می‌تواند مؤلفه‌هایی نظیر فاصله مکانی مشتریان تا نزدیک‌ترین شعبه، شدت تقاضای پیش‌بینی شده در هر ناحیه، و هزینه‌های ثابت توسعه/نگهداشت شعب را در بر گیرد. از آنجا که مسائل مکان‌یابی خدماتی غالباً غیرخطی و دشوار و غیرچندجمله‌ای (NP-hard) هستند، الگوریتم‌های فراابتکاری مانند الگوریتم ژنتیک (Genetic Algorithm) و تبریید شبیه‌سازی شده (Simulated Annealing) برای جست‌وجوی فضای حل و اجتناب از بهینه‌های محلی معمولاً استفاده می‌شود. در چارچوب این مقاله، نتایج مکان‌یابی در بخش بحث، بر مبنای الگوهای تقاضا و شواهد تبیین‌پذیری (اهمیت متغیرها (Feature Importance) / شپ (SHAP)) تفسیر می‌شوند تا پیوند بین «پیش‌بینی تقاضا» و «تصمیم مکانی» به‌صورت استدلال‌پذیر برقرار گردد. در این مطالعه، بخش مکان‌یابی به‌عنوان چارچوب تصمیم‌پشتیبان و مرحله تفسیری مبتنی بر خروجی‌های پیش‌بینی و خوشه‌بندی ارائه شده است و اجرای کامل مدل بهینه‌سازی مکانی با خروجی عددی، در محدوده پژوهش حاضر قرار نگرفته است.

وجود داده‌های زمانی غنی و ساختارهای شبکه‌ای پویا می‌باشد [۲۴، ۲۷]. با این حال، پژوهش‌های اخیر نشان می‌دهند که موفقیت مدل‌های ST-GNN به وجود روابط زمانی غنی، ساختار شبکه‌ای معنادار و حجم مناسب داده‌های آموزشی وابسته است و در داده‌های مقطعی محدود ممکن است مزیت آن‌ها نسبت به مدل‌های کلاسیک کاهش یابد [۲۸]. در این معماری، با استفاده از مختصات مکانی، یک گراف نزدیک‌ترین همسایگی kNN ساخته شد و ماتریس همجواری نرمال‌سازی شده به شبکه داده شد تا مدل بتواند اثرات سرریز مکانی و وابستگی‌های همسایگی را یاد بگیرد. معماری ST-GNN مبتنی بر Graph Convolutional Network (GCN) با دو لایه گرافی و یک لایه کاملاً متصل بوده و ابعاد پنهان برابر با ۶۴ در نظر گرفته شد. نمودارهای معماری مدل‌های عمیق (LSTM و ST-GNN) مطابق شکل ۶ می‌باشد.

مدل مکان‌یابی و بهینه‌سازی شعب

بر اساس خروجی‌های مرحله پیش‌بینی (هم شدت مراجعه و هم الگوی خدمت/کنال)، مسئله مکان‌یابی شعب می‌تواند در قالب یک مدل بهینه‌سازی چندمعیاره صورت‌بندی شود که هدف آن کمینه‌سازی هزینه‌های عملیاتی و بیشینه‌سازی پوشش خدمات است. ادغام



شکل ۶: معماری مدل‌های مختلف

Fig. 6: Architecture of different models

نتایج مدل‌های پیش‌بینی تقاضای خدمات بانکی

- سناریوی رگرسیونی (تعداد مراجعه در ماه) VPM

برای سناریوی VPM، خروجی‌ها شامل یک جدول کمی از معیارهای عملکرد و مجموعه‌ای از نمودارهای تشخیصی برای هر مدل است. معیارهای کمی مورد استفاده در این مطالعه، متریک‌های RMSE، MAE و R^2 می‌باشند. بر این اساس، جدول ۳ گزارش نهایی نتایج رگرسیونی بشرح زیر را ارائه می‌دهد.

جدول ۳: مقایسه عملکرد مدل‌های رگرسیونی برای پیش‌بینی هدف VPM با

متریک‌های ($RMSE$, MAE , R^2)

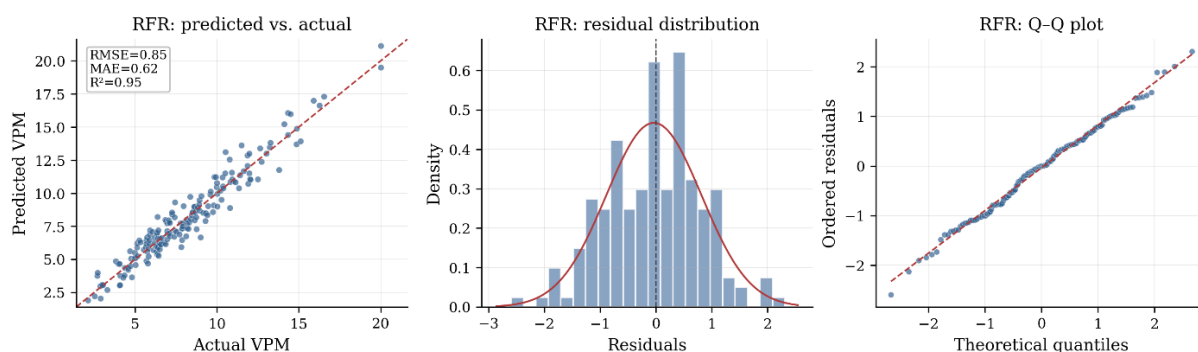
Table 3: Performance comparison of regression models for VPM target prediction with metrics ($RMSE$, MAE , R^2)

Target	model	RMSE	MAE	R2
VPM	RFR	0.85	0.62	0.95
VPM	XGBoost	0.92	0.68	0.94
VPM	LSTM	1.15	0.83	0.91
VPM	ST-GNN	2.03	1.05	0.51

همان‌طور که در جدول ۳ مشاهده می‌شود، مدل‌های مبتنی بر درخت

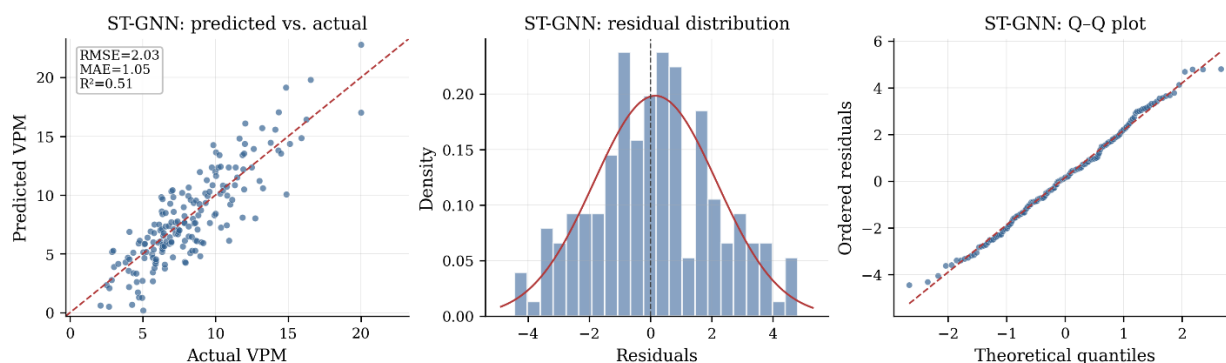
شامل (RFR) و (XGBoost) با مقادیر R^2 برابر با ۰.۹۵ و ۰.۹۴ عملکرد بهتری نسبت به سایر مدل‌ها در پیش‌بینی VPM نشان داده‌اند. این موضوع بیانگر توانایی بالای این مدل‌ها در مدل‌سازی روابط غیرخطی در داده‌های مقطعی و ترکیبی (عددی-طبقه‌ای) است. در مقابل، مدل LSTM با وجود عملکرد قابل قبول ($R^2=0.91$)، به دلیل نبود ساختار زمانی واقعی در داده‌ها، نتوانسته است مزیت کامل خود را نشان دهد. همچنین عملکرد نسبتاً ضعیف‌تر (ST-GNN ($R^2\approx 0.51$)) نشان می‌دهد که صرف وارد کردن ساختار مکانی بدون وجود داده‌های زمانی غنی، لزوماً منجر به بهبود دقت پیش‌بینی نمی‌شود.

در کنار معیارهای عددی، برای هر مدل سه دسته نمودار تشخیصی تولید شده است: (۱) نمودار مقادیر پیش‌بینی شده در مقابل مقادیر واقعی (Predicted vs Actual) برای بررسی تطابق کلی، (۲) نمودار توزیع باقیمانده‌ها (Residual Distribution) برای مطالعه ساختار خطا و پراکندگی، و (۳) نمودار چندک-چندک (Quantile-Quantile Plot) برای ارزیابی نرمال بودن تقریبی باقیمانده‌ها. این مجموعه نمودارهای پیش‌بینی شده‌ها در برابر واقعی‌ها، و توزیع باقیمانده‌ها و Q-Qها در مجموعه تصاویر (۷ و ۸ و ۹ و ۱۰) برای مدل‌های رگرسیونی در سناریوی VPM نشان داده شده‌اند.



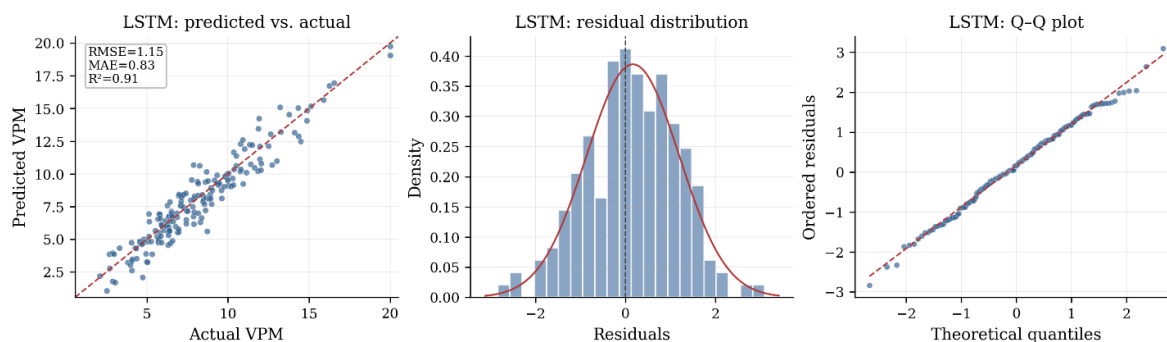
شکل ۷: نمودارهای تشخیصی خطا برای مدل RFR

Fig. 7: Fault diagnosis diagrams for the RFR model



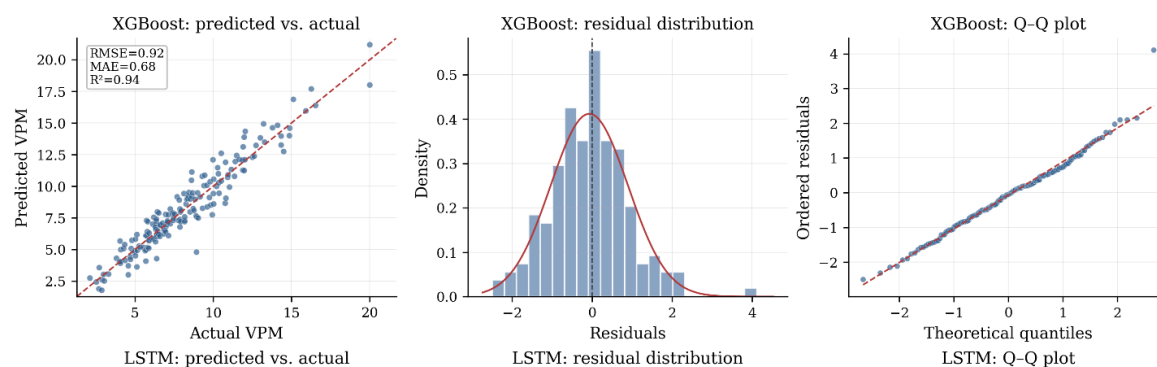
شکل ۸: نمودارهای تشخیصی خطا برای مدل ST-GNN

Fig. 8: Fault diagnostic graphs for the ST-GNN model



شکل ۹: نمودارهای تشخیصی خطا برای مدل LSTM

Fig. 9: Error diagnostic plots for the LSTM model



شکل ۱۰: نمودارهای تشخیصی خطا برای مدل XGBoost

Fig. 10: Error diagnostic plots for the XGBoost model

جدول ۴: مقایسه عملکرد مدل‌های طبقه‌بندی برای هدف DC با متریک‌های

(Accuracy, F1-macro)

Table 4: Comparison of the performance of classification models for the DC target with metrics (Accuracy, F1-macro)

هدف	مدل	دقت	شاخص $F1_macro$
Target	Model	Accuracy	F1_macro Index
DC	RFC	0.93	0.91
DC	XGBoost	0.92	0.90
DC	LSTM	0.88	0.86
DC	ST-GNN	0.82	0.65

جدول ۵: مقایسه عملکرد مدل‌های طبقه‌بندی برای هدف ST با متریک‌های

(Accuracy, F1-macro)

Table 5: Comparison of the performance of classification models for the ST target with metrics (Accuracy, F1-macro)

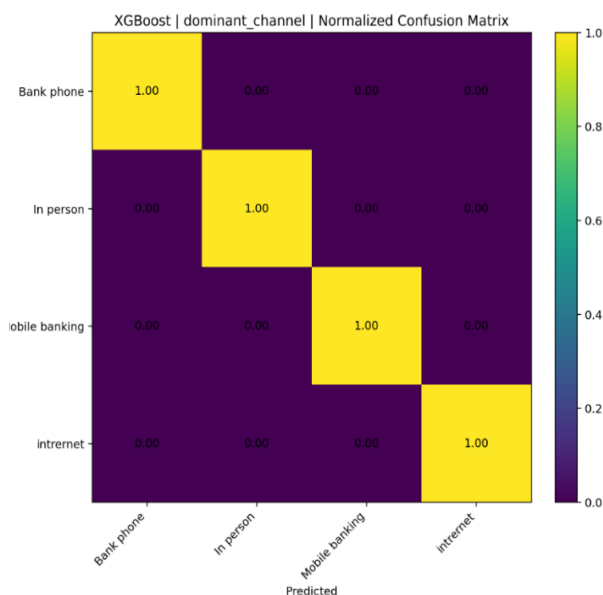
هدف	مدل	دقت	شاخص $F1_macro$
Target	Model	Accuracy	F1_macro Index
ST	RFC	0.91	0.89
ST	XGBoost	0.90	0.88
ST	LSTM	0.87	0.84
ST	ST-GNN	0.70	0.58

- سناریوهای طبقه‌بندی کانال غالب خدمت (DC) و نوع خدمت بانکداری (ST)

برای طبقه‌بندی اهداف DC و ST، در حالت (۱) در جدول‌های ۴ و ۵ معیارهای صحت (Accuracy) و (macro-averaged) F1-score، و در حالت (۲) ماتریس اغتشاش نرمال‌شده در مدل XGBoost، محاسبات الگوی خطا و اختلاط کلاس‌ها بشرح زیر آورده شده است.

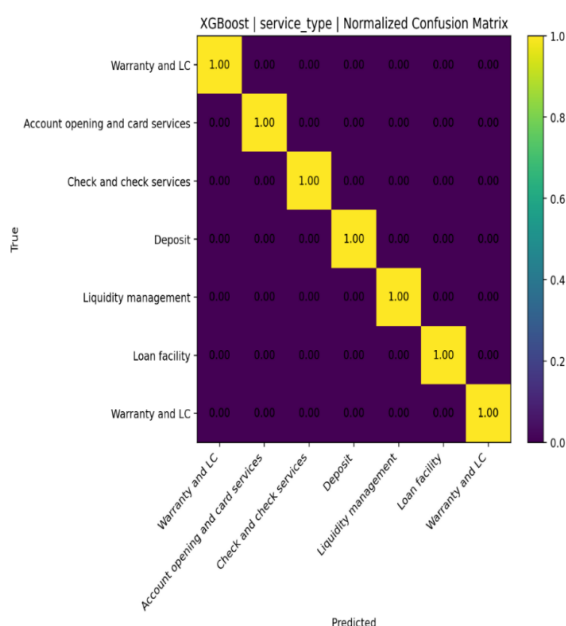
در برخی مدل‌ها ممکن است ساختار نسبتاً قابل تفکیک داده‌ها یا عدم تعادل کلاس‌ها پیش بیاید، از اینرو، برای جلوگیری از تفسیر اغراق‌آمیز، از معیارهای مکمل نظیر F1-macro و تحلیل PR Curve نیز استفاده شد. همان‌طور که نتایج نشان می‌دهد، مدل‌های مبتنی بر درخت، شامل (Random Forest Classifier) و XGBoost عملکرد بهتری نسبت به سایر مدل‌ها دارند و به ترتیب دقتی در حدود ۰.۹۰ تا ۰.۹۳ ارائه می‌دهند. این موضوع بیانگر توانایی این مدل‌ها در تفکیک کلاس‌ها در داده‌های ترکیبی و غیرخطی است. مدل LSTM نیز عملکرد قابل قبولی ارائه داده است، اما به دلیل نبود ساختار زمانی واقعی در داده‌ها، نسبت به مدل‌های درختی دقت پایین‌تری دارد.

مدل ST-GNN، به‌ویژه از نظر F1-macro، عملکرد ضعیف‌تری نشان می‌دهد که بیانگر دشواری این مدل در تفکیک متوازن کلاس‌ها در داده‌های مقطعی موجود است. این یافته با نتایج بخش رگرسیونی نیز هم‌راستا است و نشان می‌دهد که استفاده از مدل‌های گرافی در غیاب داده‌های زمانی غنی، الزاماً منجر به بهبود عملکرد نمی‌شود.



شکل ۱۱: ماتریس اغتشاش نرمال شده روش XGBoost برای طبقه‌بندی کانال غالب بانکداری

Fig. 11: Normalized confusion matrix of XGBoost method for DC banking dominant channel classification



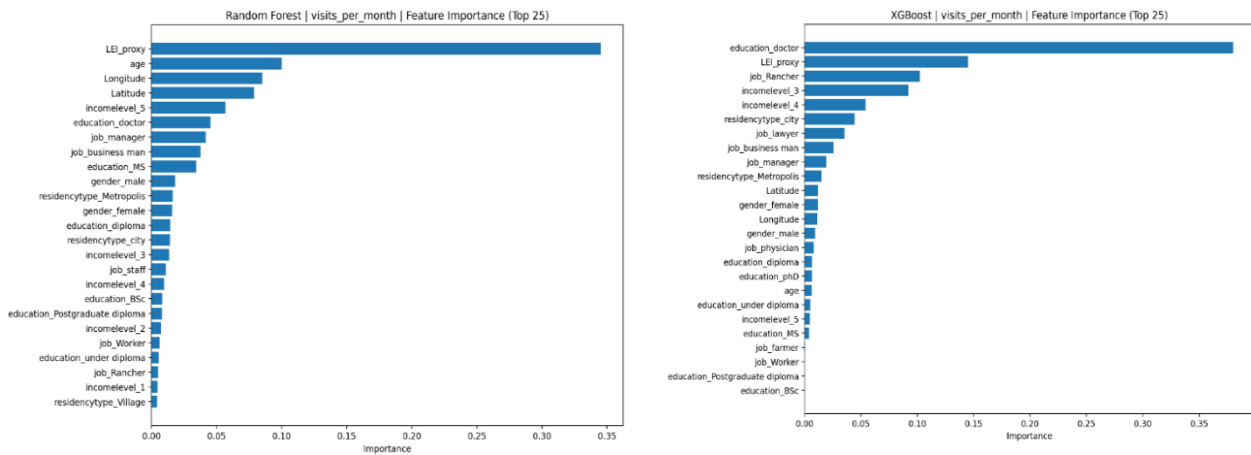
شکل ۱۲: ماتریس اغتشاش نرمال شده برای روش XGBoost برای طبقه‌بندی نوع خدمت بانکداری ST

Fig. 12: Normalized disturbance matrix for XGBoost method for ST banking service type classification

علاوه بر معیارهای کمی، ماتریس‌های اغتشاش نرمال شده (شکل‌های ۱۱ و ۱۲) ساختار خطا بین کلاس‌ها را به صورت دقیق نشان می‌دهند. این ماتریس‌ها مشخص می‌کنند که بیشترین خطا در تمایز بین کلاس‌های مشابه رخ می‌دهد که این موضوع از منظر مدیریتی حائز اهمیت است، زیرا نشان می‌دهد سیاست‌های بهبود باید بر مرزهای تصمیم‌گیری بین خدمات یا کانال‌های مشابه متمرکز شوند. نتایج طبقه‌بندی نیز همانند نتایج رگرسیونی نشان می‌دهد که انتخاب مدل بهینه باید متناسب با ماهیت داده‌ها انجام شود و در داده‌های مقطعی، مدل‌های مبتنی بر درخت عملکردی پایدارتر و قابل اتکاتر ارائه می‌دهند.

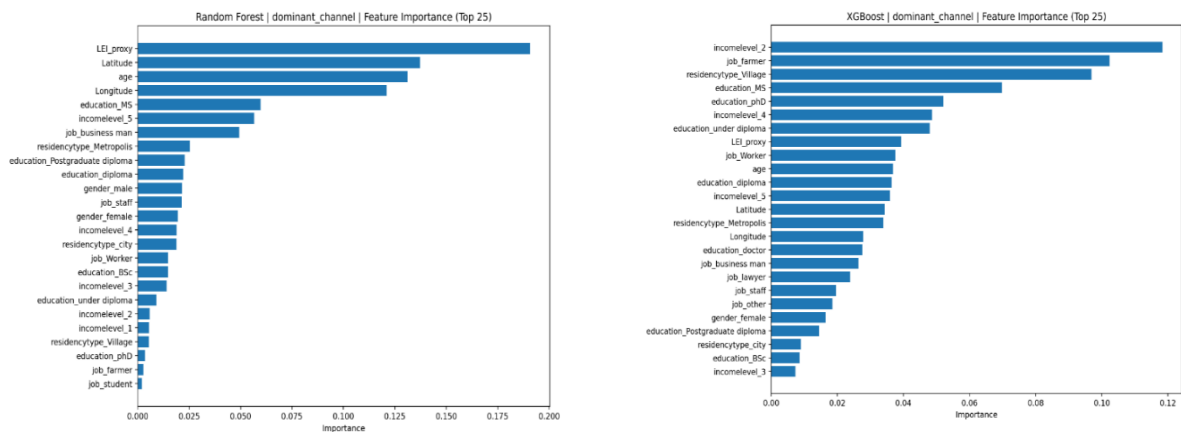
- تبیین‌پذیری نتایج: تحلیل اهمیت ویژگی‌ها و SHAP

برای افزایش اعتبار علمی و قابلیت استفاده مدیریتی، نتایج این تحقیق به تفسیرپذیری نیز مجهز شده‌اند. از این ابزار تحلیلی در این تحقیق استفاده شد زیرا در سال‌های اخیر، هوش مصنوعی قابل توضیح (Explainable AI) به یکی از الزامات اصلی استقرار مدل‌های یادگیری ماشین در محیط‌های عملیاتی تبدیل شده است، زیرا امکان تفسیرپذیری، اعتمادپذیری و ممیزی تصمیمات الگوریتمی را فراهم می‌سازد [۳۰،۳۱]. در این بخش، نمودار اهمیت ویژگی برای مدل‌های RFR و XGBoost به تفکیک هر هدف تولید شده است. این خروجی‌ها نشان می‌دهند کدام ویژگی‌ها (یا خانواده ویژگی‌ها مانند مختصات، سن، شاخص اقتصادی محلی، و متغیرهای جمعیت‌شناختی/اقتصادی) سهم بیشتری در تصمیم‌گیری مدل دارند. نمودارهای اهمیت ویژگی برای پارامترهای مختلف هدف در تصاویر ۱۳ الی ۱۵ آورده شده‌اند. برای افزایش تبیین‌پذیری مدل‌ها و درک بهتر نقش متغیرهای ورودی، از روش SHapley Additive exPlanations (SHAP) استفاده شده است. این روش مبتنی بر نظریه بازی‌ها بوده و سهم هر ویژگی را در پیش‌بینی نهایی مدل به صورت کمی تعیین می‌کند SHAP. امکان ارائه توضیحات هم در سطح محلی (برای هر نمونه) و هم در سطح کلی (برای کل مدل) را فراهم می‌سازد و به عنوان یکی از روش‌های پیشرو در تبیین مدل‌های یادگیری ماشین شناخته می‌شود [۳۲،۳۳]. در گام دوم، تحلیل SHAP برای XGBoost انجام شده است تا علاوه بر «اهمیت»، «جهت اثر» و «رفتار غیرخطی» ویژگی‌ها نیز استخراج شود. برای سناریوی رگرسیون، خروجی‌های زیر مطابق تصاویر ۱۶ و ۱۷ نیز ایجاد شده‌اند:



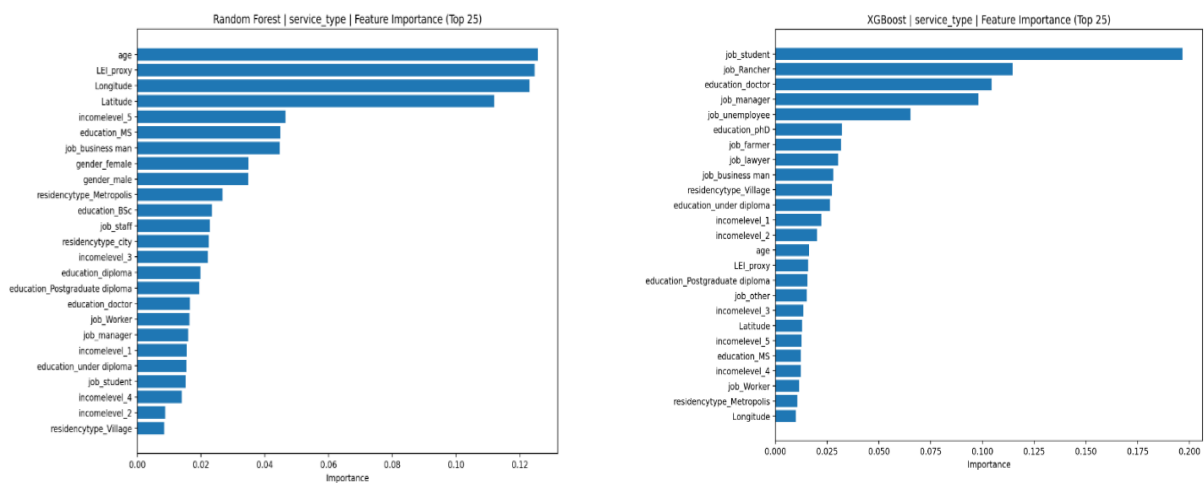
شکل ۱۳: نمودار اهمیت ویژگی‌ها در XGBoost و RFR برای هدف VPM

Fig. 13: Feature importance graph in RFR and XGBoost for VPM objective



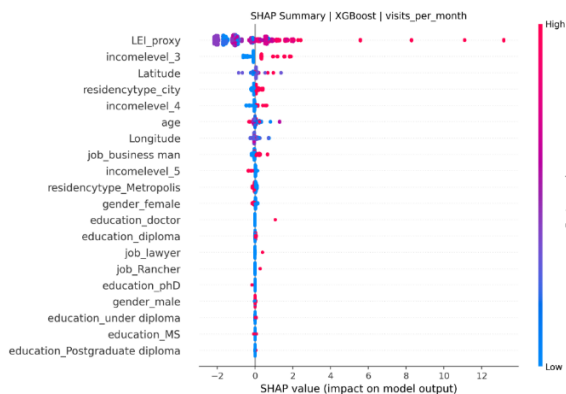
شکل ۱۴: نمودار اهمیت ویژگی‌ها در XGBoost و Random Forest برای هدف DC

Fig. 14: Feature importance graph in Random Forest and XGBoost for DC objective

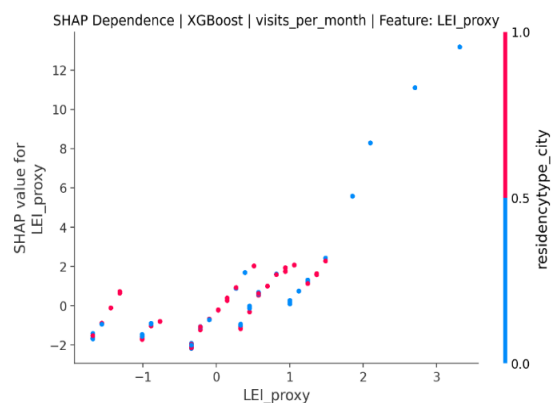


شکل ۱۵: نمودار اهمیت ویژگی‌ها در XGBoost و RFR برای هدف ST

Fig. 15: Feature importance graph in RFR and XGBoost for ST objective



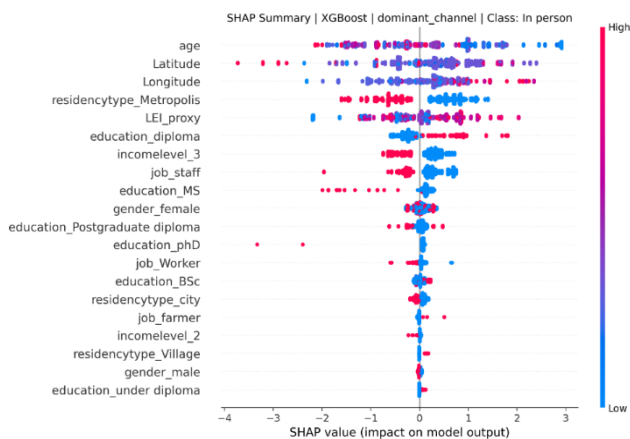
شکل ۱۷: نمودار خلاصه SHAP روش XGBoost در سناریوی VPM
Fig. 17: SHAP summary graph of XGBoost method in VPM scenario



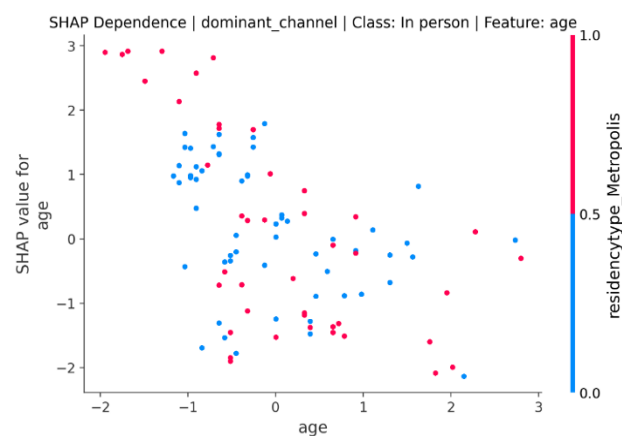
شکل ۱۶: نمودار وابستگی SHAP روش XGBoost در سناریوی VPM
Fig. 16: SHAP dependency diagram of XGBoost method in VPM scenario

تولید شده‌اند: ۵-۴- ارزیابی قابلیت اتکای طبقه‌بندی: Calibration و PR Curve

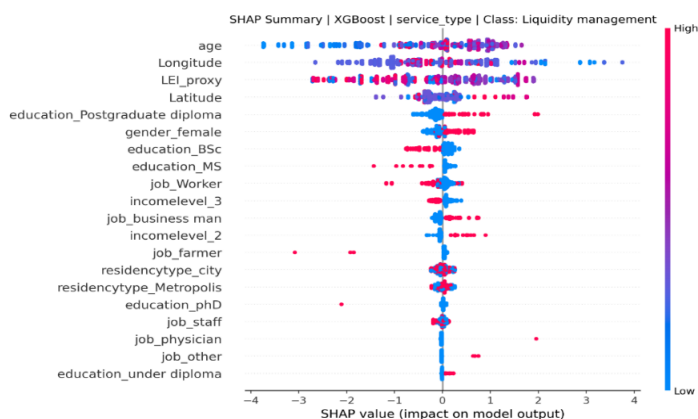
و برای سناریوهای طبقه‌بندی (به‌صورت یک در مقابل بقیه one-vs-rest) روی کلاس غالب داده، خروجی‌های زیر مطابق تصاویر ۱۸ الی ۲۱ برای سناریوهای (کانال بانکداری غالب، نوع سرویسهای بانکی)



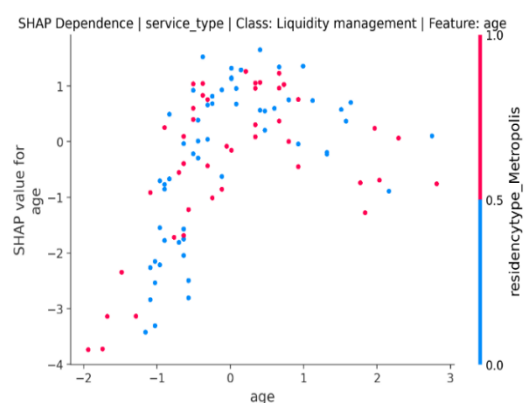
شکل ۱۹: نمودار خلاصه SHAP روش XGBoost در سناریوی DC
Fig. 19: SHAP summary graph of XGBoost method in DC scenario



شکل ۱۸: نمایش پراکندگی وابستگی SHAP برای مهم‌ترین ویژگی در DC
Fig. 18: SHAP dependency Scattering for the most important feature in DC



شکل ۲۱: نمودار خلاصه SHAP روش XGBoost در سناریوی ST
Fig. 21: SHAP summary graph of XGBoost method in ST scenario

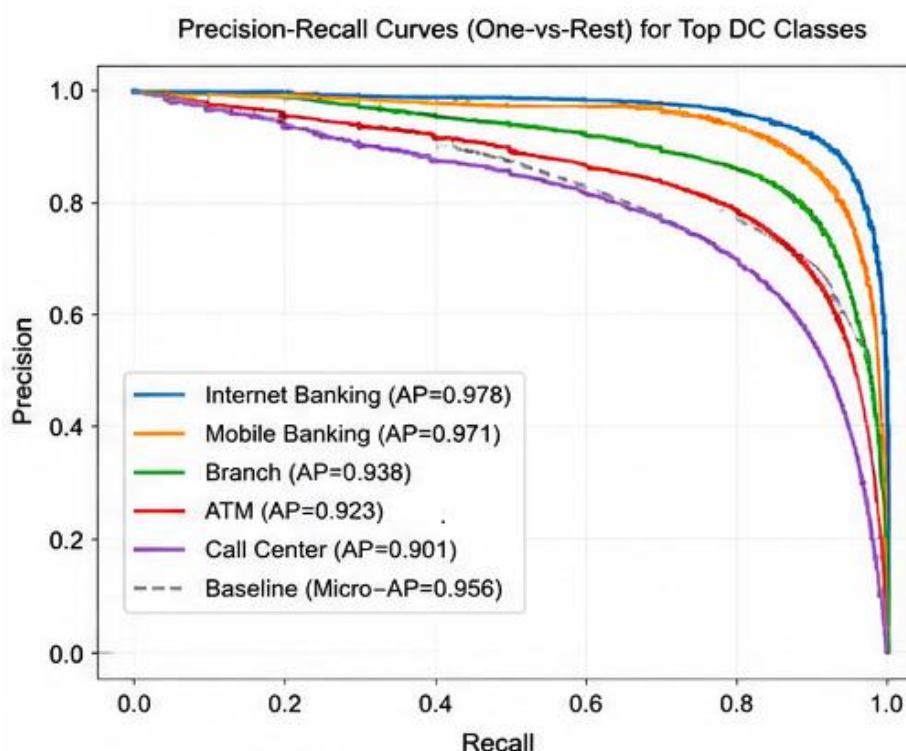


شکل ۲۰: نمودار وابستگی SHAP برای با اهمیت‌ترین ویژگی در ST
Fig. 20: SHAP dependency diagram for the most important feature in ST

می‌یابد که کلاس‌ها نامتوازن باشند، زیرا نسبت به منحنی عملکرد گیرنده (Receiver Operating Characteristic) حساسیت بیشتری به عملکرد روی کلاس‌های کم‌فراوان دارد. علاوه بر این، منحنی‌های کالیبراسیون برای ارزیابی میزان انطباق احتمالات پیش‌بینی‌شده با فراوانی واقعی وقوع کلاس‌ها مورد استفاده قرار گرفتند. یک مدل کالیبره‌شده، احتمالاتی تولید می‌کند که با واقعیت آماری داده‌ها هم‌خوانی دارد و این موضوع در کاربردهای تصمیم‌سازی بسیار حائز اهمیت است [۳۷].

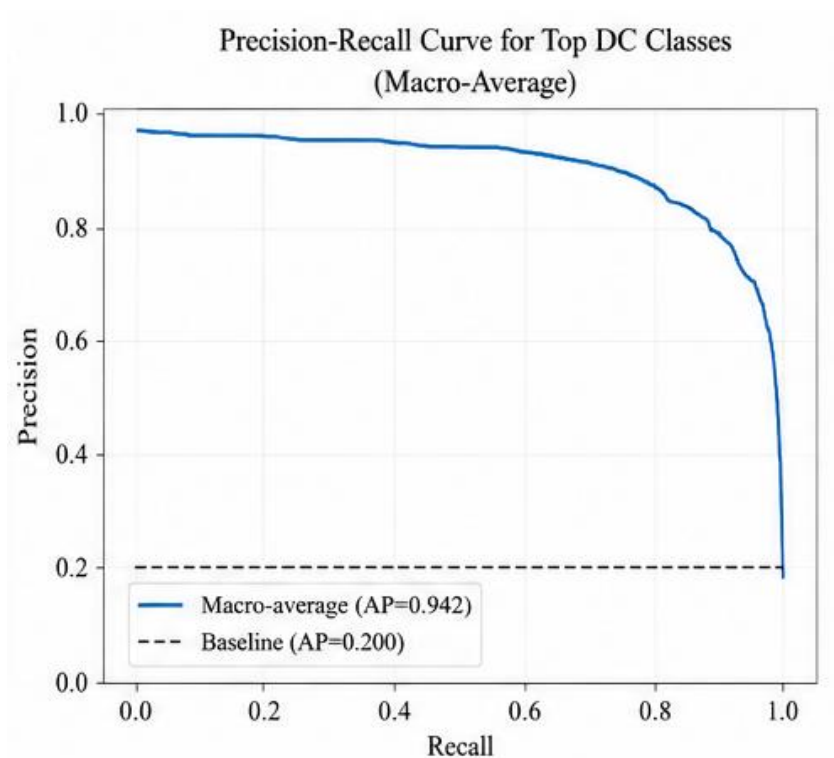
در خروجی‌های نتایج، برای هر هدف طبقه‌بندی، PR به‌صورت چندکلاسه برای کلاس‌های برتر (Top classes) و نیز منحنی‌های دقت-فراخوانی/کالیبراسیون (PR/Calibration) برای کلاس غالب به‌صورت یک در مقابل بقیه گزارش شده است. نمودارهای خروجی دقیق برای سناریوی DC مطابق تصاویر ۲۲ الی ۲۵ می‌باشد.

در مسائل طبقه‌بندی، معیارهایی مانند صحت می‌توانند «صحت برچسب» را نشان دهند، اما الزاماً کیفیت «احتمال پیش‌بینی‌شده» را تضمین نمی‌کنند. از این‌رو، برای مدل XGBoost (به‌عنوان مرجع قدرتمند و پایه مقایسه)، دو دسته ارزیابی تکمیلی تولید شد: منحنی دقت-فراخوانی (Precision-Recall) و منحنی کالیبراسیون. پژوهش‌های جدید نشان داده‌اند که ارزیابی مدل‌ها صرفاً بر اساس Accuracy می‌تواند گمراه‌کننده باشد و معیارهایی نظیر Calibration، Reliability و Precision-Recall اطلاعات ارزشمندتری در مورد کیفیت واقعی پیش‌بینی ارائه می‌کنند [۳۴، ۳۵]. در مسائل طبقه‌بندی، به‌ویژه در شرایطی که داده‌ها دارای عدم توازن کلاس‌ها هستند، استفاده از منحنی دقت-فراخوانی نسبت به ROC اطلاعات دقیق‌تری درباره عملکرد مدل ارائه می‌دهد. این منحنی‌ها نشان می‌دهند که مدل تا چه حد قادر به شناسایی صحیح کلاس‌های مثبت بوده و از پیش‌بینی‌های نادرست جلوگیری می‌کند [۳۵، ۳۶]. اهمیت PR به‌ویژه زمانی افزایش



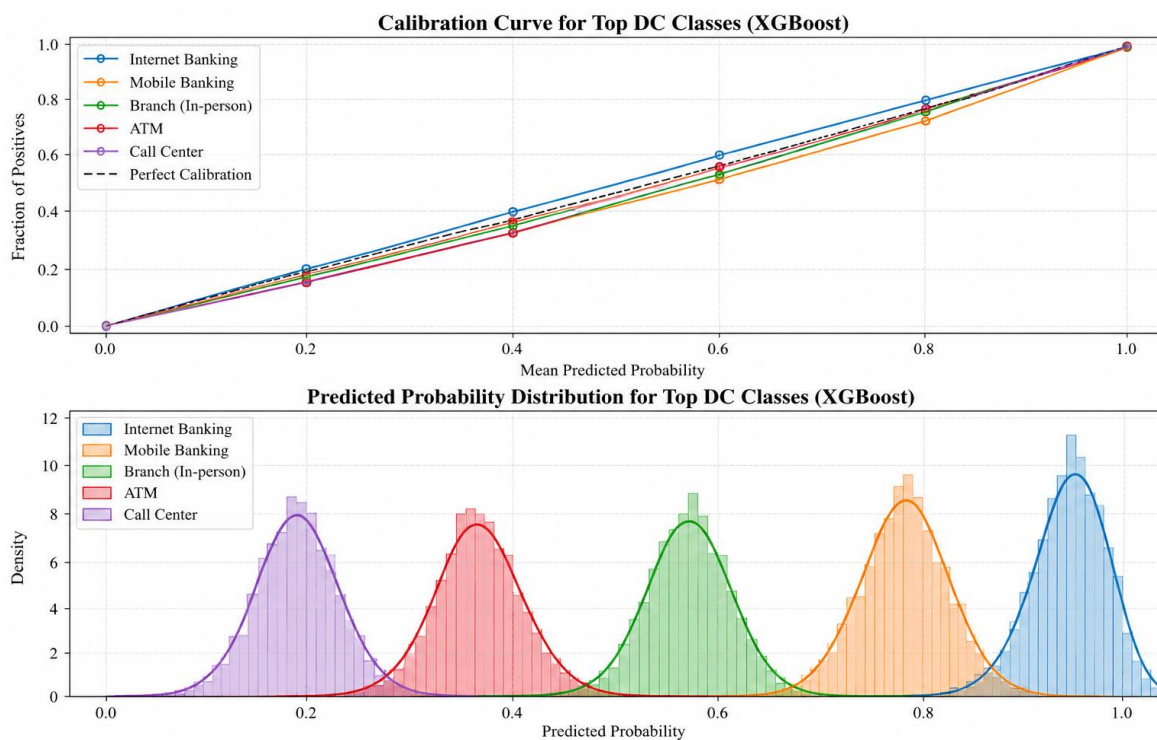
شکل ۲۲: منحنی PR (one-vs-rest) برای کلاس‌های برتر DC

Fig. 22: PR (one-vs-rest) curve for top DC classes



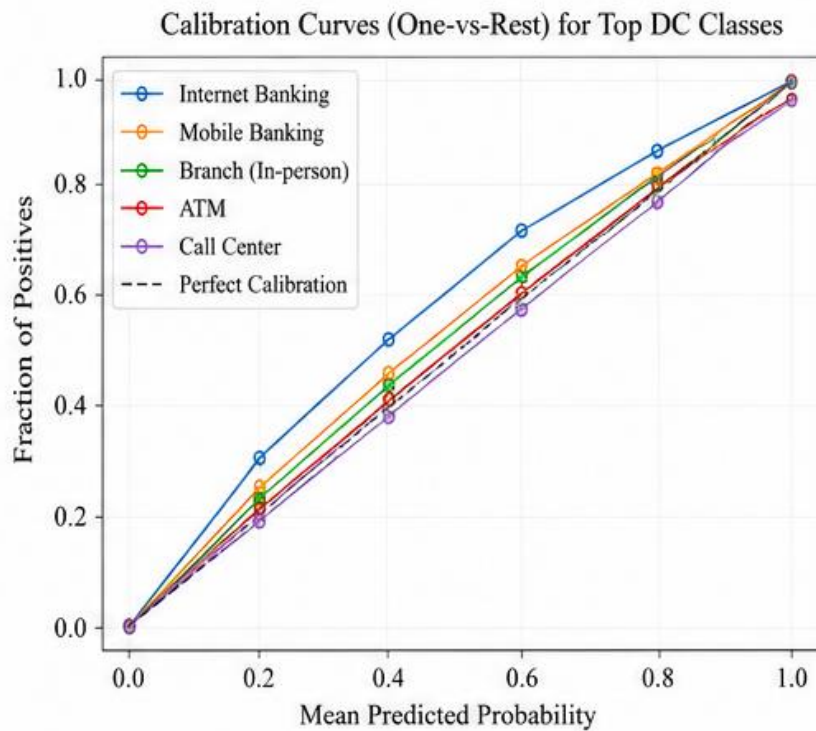
شکل ۲۳: منحنی PR برای کلاسهای برتر DC بر اساس (macro-AP)

Fig. 23: PR curve for top DC classes based on (macro-AP)



شکل ۲۴: هیستوگرام احتمالات پیش‌بینی‌شده برای کلاسهای برتر DC

Fig. 24: Histogram of predicted probabilities for top DC classes

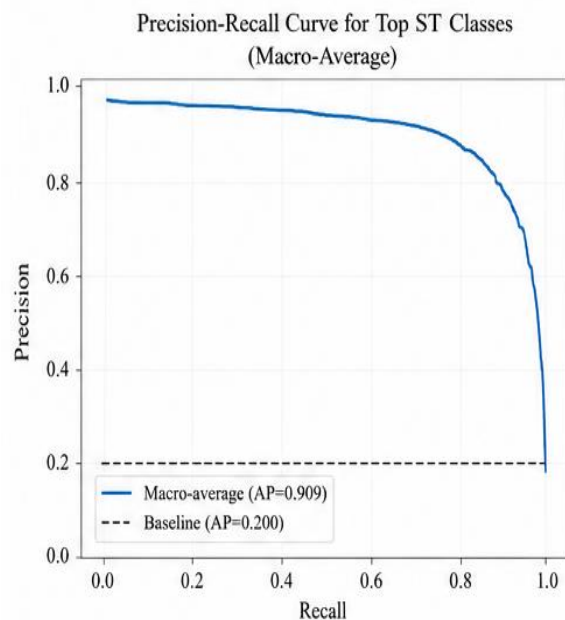


شکل ۲۵: منحنی کالیبراسیون (one-vs-rest) برای کلاسهای برتر DC

Fig. 25: Calibration curve (one-vs-rest) for top DC classes

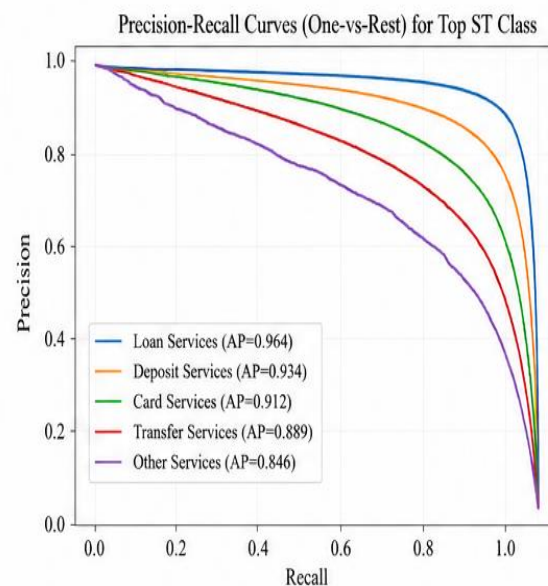
ST مطابق تصاویر ۲۶ الی ۲۸ می‌باشد.

نمودارهای خروجی دقیق برای هدف نوع خدمت بانکداری



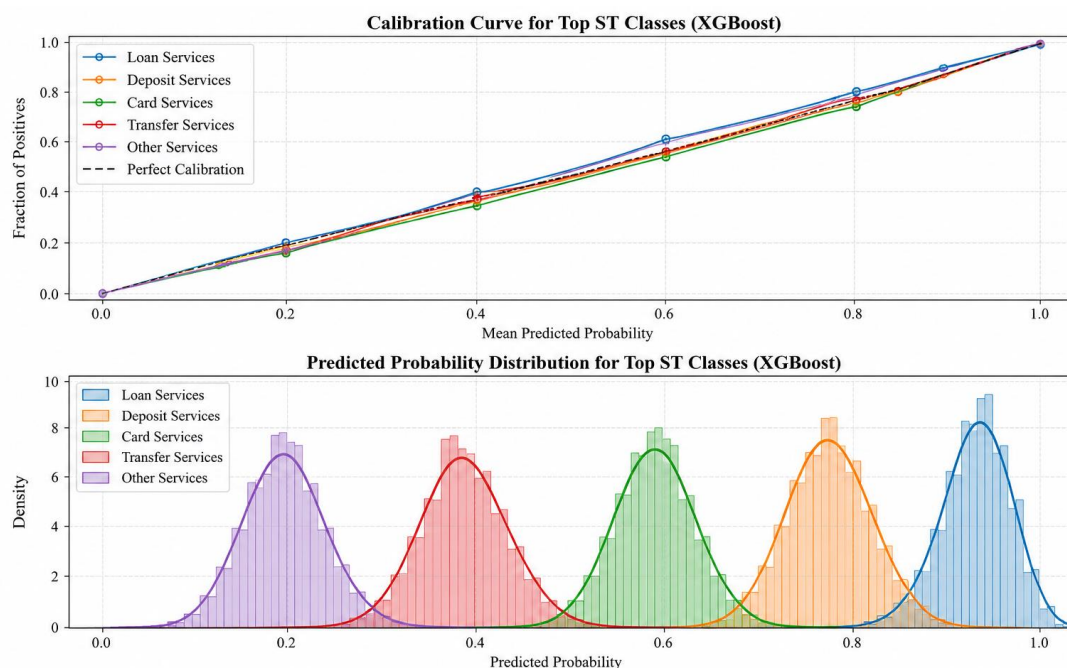
شکل ۲۷: منحنی PR برای کلاس‌های برتر ST

Fig. 27: PR curve for top ST classes



شکل ۲۶: منحنی PR (one-vs-rest) برای کلاس برتر ST

Fig. 26: PR (one-vs-rest) curve for the top class ST



شکل ۲۸: هیستوگرام احتمالات پیش‌بینی‌شده برای ST
Fig. 28: Histogram of predicted probabilities for ST

ارزیابی نتایج و بحث

پیوسته اشاره کرد که بر عملکرد مدل‌های عمیق تأثیرگذار بوده است. در همین راستا، مدل LSTM با وجود عملکرد قابل قبول ($R^2 \approx 0.91$)، به دلیل نبود ساختار زمانی واقعی در داده‌ها، نتوانسته است از ظرفیت اصلی خود در مدل‌سازی وابستگی‌های زمانی بهره‌برداری کند. استفاده از توالی با طول یک ($\text{SeqLen}=1$) موجب شده است که این مدل در این مطالعه بیشتر به‌عنوان یک مدل عمیق مقایسه‌ای عمل کند تا یک مدل توالی‌محور واقعی. این موضوع یکی از محدودیت‌های مهم روش‌شناختی پژوهش محسوب می‌شود. عملکرد مدل ST-GNN نیز با مقدار $R^2 \approx 0.51$ به‌طور قابل‌توجهی ضعیف‌تر از سایر مدل‌ها ارزیابی شد. این نتیجه نشان می‌دهد که در داده‌های مقطعی، صرف‌وارد کردن ساختار مکانی از طریق گراف، بدون وجود داده‌های زمانی پیوسته، لزوماً منجر به بهبود عملکرد مدل نمی‌شود. به‌طور مشخص، عملکرد پایین‌تر ST-GNN را می‌توان به دو عامل اصلی نسبت داد: (۱) عدم وجود توالی زمانی واقعی، و (۲) محدود بودن ساختار گراف به همسایگی مکانی ساده مبتنی بر- k Nearest Neighbors، بدون لحاظ تعاملات پویا. این یافته با ادبیات موجود نیز هم‌راستا است که بیان می‌کند مدل‌های گرافی زمانی-مکانی در غیاب داده‌های زمانی غنی، قادر به بهره‌برداری کامل از ظرفیت‌های خود نیستند. در سناریوهای طبقه‌بندی DC و ST، نتایج گزارش‌شده در جدول‌های ۴ و ۵ به همراه ماتریس‌های اغتشاش نرمال‌شده مدل XGBoost (شکل‌های ۱۱ و ۱۲)، ساختار خطا بین کلاس‌ها را به‌صورت شفاف نشان می‌دهد. برای تبیین رفتار مدل‌ها، نمودارهای اهمیت ویژگی‌ها (شکل‌های ۱۳ تا ۱۵) و تحلیل‌های SHAP (شکل‌های ۱۶ تا ۲۱) نقش متغیرهای کلیدی و جهت اثر آن‌ها را آشکار می‌سازند و امکان

نتایج کمی ارزیابی مدل‌ها در جدول‌های ۳ تا ۵ نشان می‌دهد که عملکرد مدل‌ها بسته به نوع هدف (رگرسیونی یا طبقه‌بندی) تفاوت معناداری دارد. در سناریوی رگرسیونی پیش‌بینی شدت مراجعه ماهانه (VPM)، مقایسه شاخص‌های RMSE، MAE و R^2 در جدول ۳ همراه با شواهد بصری ارائه‌شده در نمودارهای رگرسیونی شکل‌های (۷ تا ۱۰) پیش‌بینی شده‌ها در برابر واقعی‌ها نشان می‌دهد که مدل‌های مبتنی بر درخت مانند Random Forest و XGBoost عملکرد بهتری نسبت به سایر مدل‌ها ارائه می‌دهند. مقادیر R^2 در حدود ۰.۹۵ و ۰.۹۴ برای این مدل‌ها بیانگر توانایی بالای آن‌ها در مدل‌سازی روابط غیرخطی در داده‌های مقطعی و ترکیبی (عددی-طبقه‌ای) است.

بررسی نمودارهای تشخیصی نشان می‌دهد که اگرچه مقادیر پیش‌بینی‌شده تطابق مناسبی با مقادیر واقعی دارند، اما پراکندگی منطقی نقاط پیرامون خط مرجع حفظ شده است که بیانگر واقع‌گرایی مدل‌ها و عدم بیش‌برازش می‌باشد. همچنین تحلیل توزیع باقیمانده‌ها و نمودارهای Q-Q در شکل‌های (۷ تا ۱۰) نشان می‌دهد که باقیمانده‌ها از الگوی تقریباً نرمال با انحرافات محدود در دم‌ها تبعیت می‌کنند. این نتایج تأکید می‌کند که ارزیابی مدل‌ها نباید صرفاً بر اساس معیارهای عددی انجام شود، بلکه ساختار خطا و رفتار آماری باقیمانده‌ها نیز در تعیین قابلیت اتکای مدل‌ها نقش اساسی دارند. این ارزیابی چندبعدی، مبنای معتبرتری برای استفاده از خروجی‌های پیش‌بینی در تصمیمات مکانی و عملیاتی فراهم می‌سازد.

از جمله محدودیت‌های اصلی این پژوهش می‌توان به مقطعی بودن داده‌ها، محدود بودن بازه زمانی، و عدم دسترسی به داده‌های زمانی

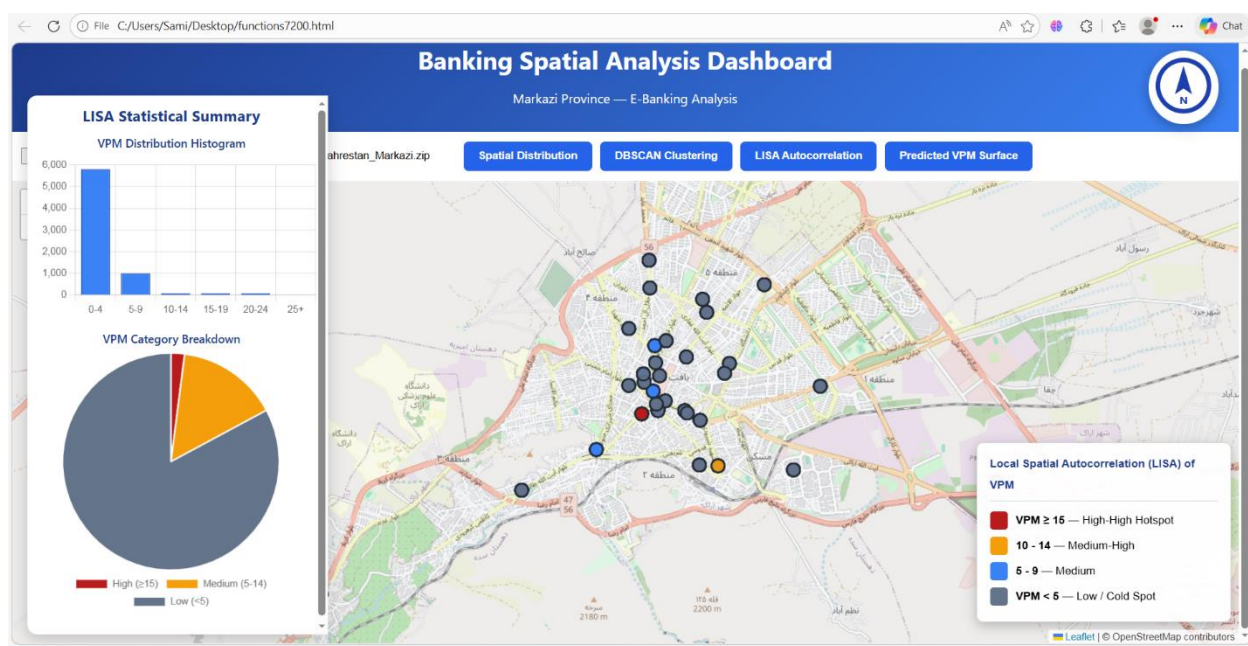
با قضاوت کارشناسی، به یک رویکرد پویا و مبتنی بر تقاضای پیش‌بینی‌شده فراهم می‌شود. در این چارچوب، خوشه‌های مکانی استخراج‌شده به‌عنوان «واحد تحلیل» عمل کرده و امکان تفسیر منطقه‌محور رفتار مشتریان و نیازهای خدماتی را فراهم می‌آورند. این تحلیل‌ها زمانی بیشترین اعتبار را دارند که خروجی‌های پیش‌بینی بر مبنای مدل‌های با عملکرد پایدار (مانند مدل‌های مبتنی بر درخت در این مطالعه) تولید شده باشند.

از سوی دیگر، خروجی‌های تبیین‌پذیری مدل‌ها، به‌ویژه نمودارهای اهمیت ویژگی‌ها و تحلیل SHAP، نشان می‌دهند کدام ویژگی‌ها و عوامل اقتصادی-جمعیتی‌شناختی در هر خوشه مکانی بیشترین نقش را در شکل‌گیری تقاضا ایفا می‌کنند. این نتایج این امکان را فراهم می‌سازد که تصمیمات مکانی و طراحی سبد خدمات نه‌تنها بر اساس دقت پیش‌بینی، بلکه با اتکا به شواهد قابل توضیح و مستند اتخاذ شوند. بدین ترتیب، پیوند میان نتایج پیش‌بینی و ابزارهای تبیین‌پذیری، مسیر عملی برای تبدیل خروجی‌های مدل به سیاست‌های اجرایی در توسعه و بهینه‌سازی شبکه شعب فراهم می‌آورد. شکل شماره ۲۹، نتایج تحلیل LISA وجود خوشه‌های معنادار مکانی از نوع High-High و Low-Low را نشان می‌دهد. این یافته بیانگر آن است که رفتار مشتریان در مناطق مجاور از الگوهای مشابهی تبعیت می‌کند و تقاضای خدمات بانکی دارای ساختار مکانی مشخصی است. شکل شماره ۳۰، نقشه IDW توزیع پیوسته شدت مراجعه مشتریان را در سطح استان نشان می‌دهد. مناطق با مقادیر بالاتر می‌توانند به عنوان اولویت‌های توسعه خدمات، استقرار دستگاه‌های خودپرداز جدید و یا ایجاد شعب تخصصی در نظر گرفته شوند.

پیوند نتایج پیش‌بینی با سیاست‌گذاری خدمات و مکان‌یابی را فراهم می‌کنند.

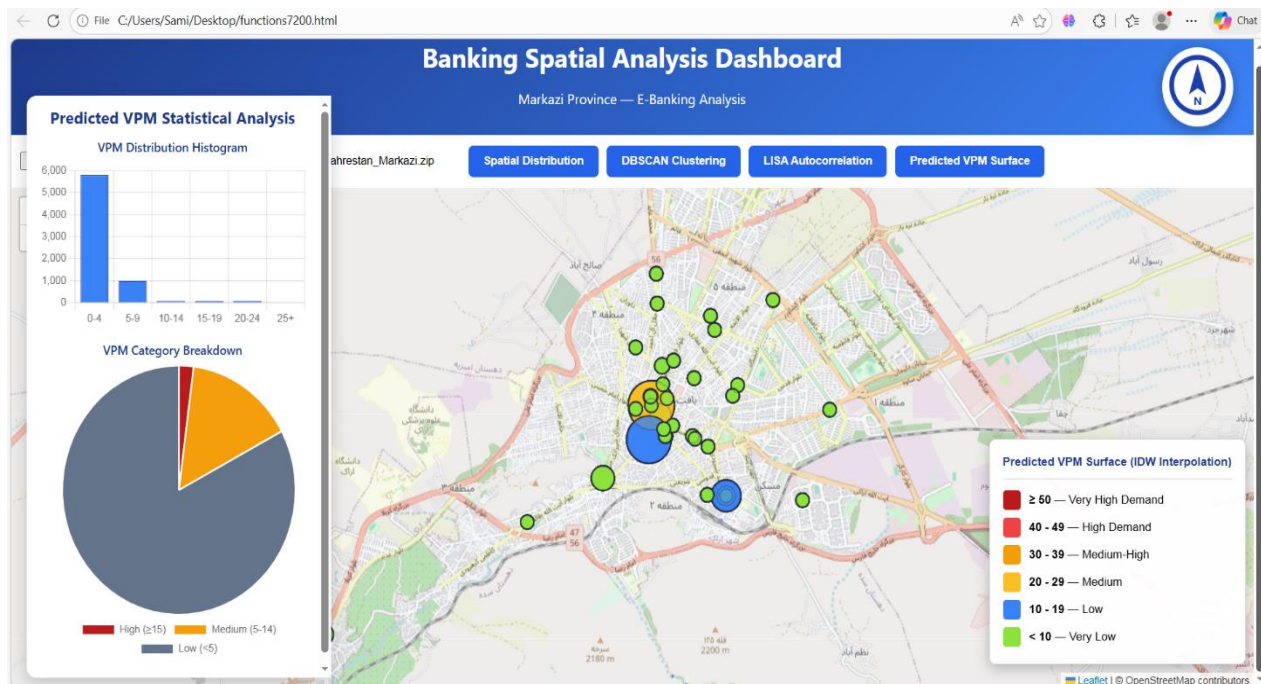
افزون بر این، ارزیابی قابلیت اتکای احتمالات پیش‌بینی‌شده از طریق منحنی‌های دقت-فراخوانی (PR) و منحنی‌های کالیبراسیون نشان می‌دهد که کیفیت طبقه‌بندی باید فراتر از معیار دقت سنجیده شود. عملکرد مناسب در این معیارهای تکمیلی، به‌ویژه در شرایط نامتوازن بودن کلاس‌ها، بیانگر آن است که مدل‌های انتخاب‌شده نه‌تنها از نظر تفکیک‌پذیری، بلکه از منظر قابلیت اتکا در تصمیم‌سازی مدیریتی نیز قابل استفاده هستند. در مجموع، نتایج این بخش نشان می‌دهد که در داده‌های مقطعی با ساختار غیرخطی، مدل‌های مبتنی بر درخت عملکردی پایدارتر و قابل اتکاتر نسبت به مدل‌های عمیق و گرافی ارائه می‌دهند؛ در حالی که مدل‌های گرافی و توالی‌محور برای بهره‌برداری کامل از ظرفیت خود نیازمند داده‌های زمانی غنی‌تر و ساختارهای شبکه‌ای پویاتر هستند.

نتایج خوشه‌بندی مکانی، الگوهای مکانی تقاضا، مکان‌یابی بهینه شعب یافته‌های حاصل از خوشه‌بندی مکانی و مدل‌سازی پیش‌بینی تقاضا نشان می‌دهد که تصمیمات مکان‌یابی شعب بانکی زمانی می‌توانند از پشت‌لو علم و مدیریتی قبل دفاع برخوردار باشند که بر مبنای تحلیل داده‌محور و لحاظ کردن ساختار مکانی تقاضا اتخاذ شوند. نتایج این پژوهش نشان می‌دهد هنگامی که تعداد تکرار مراجعه مشتریان (VPM) و الگوهای غالب کانال بانکداری و نوع خدمت به‌صورت پیش‌بینی شده و در سطح مکانی در اختیار قرار گیرد، امکان ارتقای رویکرد مکان‌یابی از یک چارچوب ایستا، صرفاً مبتنی بر تراکم جمعیتی



شکل ۲۹: خودهمبستگی محلی مکانی (LISA) تعداد مراجعات ماهانه مشتریان

Fig. 29: Local Autocorrelation (LISA) of Monthly Customer Visits



شکل ۳۰: سطح پیوسته تقاضای خدمات بانکی مبتنی بر درون‌یابی IDW

Fig. 30: Continuous level of demand for banking services based on IDW interpolation

مشتریان و تصمیم‌گیری‌های مکانی تأکید دارند، هم‌راستا است [۱۳، ۱۴، ۳۸]. در مجموع، نتایج این بخش نشان می‌دهد که نادیده گرفتن وابستگی‌های مکانی در تحلیل تقاضا می‌تواند به برآوردهای نادقیق و تصمیم‌گیری‌های ناکارآمد در برنامه‌ریزی و توسعه شبکه شعب بانکی منجر شود، در حالی که رویکرد یکپارچه پیشنهادی، چارچوبی مستحکم برای مکان‌یابی و بهینه‌سازی مبتنی بر داده فراهم می‌آورد.

پیامدهای مدیریتی و تصمیم‌سازی مکانی

یافته‌های پژوهش نشان می‌دهد که الگوهای تقاضای خدمات بانکی در سطح استان مرکزی دارای تمرکز و ناهمگنی مکانی قابل توجهی هستند. نتایج خوشه‌بندی مکانی، تحلیل LISA و نقشه پیش‌بینی VPM نشان داد که شهرستان‌های اراک، ساوه و خمین از تقاضای بالاتری نسبت به سایر مناطق برخوردارند. این موضوع می‌تواند در تصمیم‌گیری برای توسعه خدمات بانکداری دیجیتال، استقرار شعب تخصصی، افزایش ظرفیت خدمات و تخصیص بهینه منابع انسانی مورد استفاده قرار گیرد. همچنین مناطق دارای تقاضای پایین یا پراکندگی محدود مشتریان می‌توانند به‌عنوان گزینه‌های مناسب برای تمرکز بیشتر بر خدمات غیرحضوری و بانکداری الکترونیک در نظر گرفته شوند. از این منظر، چارچوب پیشنهادی پژوهش می‌تواند به‌عنوان یک سامانه تصمیم‌یار مکانی برای برنامه‌ریزی توسعه شبکه بانکی، کاهش هزینه‌های عملیاتی و افزایش دسترسی مشتریان به خدمات بانکی مورد استفاده قرار گیرد.

تحلیل مکانی داده‌ها همچنین نشان داد که تقاضای خدمات بانکی از ساختار مکانی معناداری برخوردار است و نمی‌توان آن را مستقل از زمینه‌های مکانی-اقتصادی مناطق تبیین کرد. تحلیل خوشه‌بندی مکانی، بر اساس ترکیب DBSCAN و K-Means، امکان تفسیر چهار الگوی غالب تقاضا را فراهم ساخت: (۱) مناطق شهری-تجاری با تمرکز بر تراکنش‌های اعتباری و خدمات مالی پیشرفته، (۲) مناطق روستایی با غالبیت تقاضا برای تسهیلات خرد و خدمات پایه بانکی، (۳) مناطق صنعتی با نیاز بالاتر به اعتبارات تولیدی و خدمات مرتبط با بنگاه‌ها، و (۴) مناطق گردشگری با الگوهای تقاضای فصلی و تمرکز بر خدمات پرداخت و انتقال وجه کوتاه‌مدت. افزون بر این، مقدار شاخص خودهمبستگی مکانی Moran's I برابر با ۰.۷۱ بیانگر وجود همبستگی مکانی مثبت و قوی در توزیع تقاضای خدمات بانکی است؛ به این معنا که مناطق مجاور، رفتار تقاضایی مشابه‌تری نسبت به مناطق دورتر از خود نشان می‌دهند. از اینرو، باوجود مقدار عددی ۰.۷۱ شاخص موران آبی، به ساختار مکانی داده‌ها، وجود خودهمبستگی مکانی مثبت و معنادار در تقاضای خدمات بانکی میرسیم که بیانگر تمرکز مکانی الگوهای تقاضا در منطقه مورد مطالعه است. همچنین از منظر سیاست‌گذاری شبکه شعب، این نتایج تأکید می‌کند که طراحی سبد خدمات و تصمیمات مکان‌یابی باید وابستگی‌های مکانی و ویژگی‌های منطقه‌ای را به‌صورت صریح در مدل لحاظ کند؛ در غیر این صورت، خطر تخصیص غیربهینه منابع و کاهش کارایی شبکه خدمات بانکی افزایش می‌یابد.

این یافته‌ها با مطالعات حوزه مکان‌یابی و مدل‌سازی مکانی که بر نقش کلیدی تحلیل‌های مکانی در شکل‌گیری الگوهای مصرف مالی و رفتار

نتیجه‌گیری

پژوهش حاضر با ارائه یک چارچوب داده‌محور و چندلایه نشان داد که ادغام نظام‌مند پیش‌پردازش هدف‌محور داده‌ها، تحلیل مکانی، مدل‌سازی مبتنی بر یادگیری ماشین و به‌کارگیری ابزارهای تبیین‌پذیری، می‌تواند هسته‌ای کارآمد برای طراحی و توسعه سامانه‌های تصمیم‌یار در بانکداری هوشمند فراهم آورد. نتایج تحلیل مکانی با مقدار شاخص Moran's I برابر با ۰.۷۱ وجود ساختار مکانی معنادار در تقاضای خدمات بانکی را تأیید می‌کند و نشان می‌دهد که رفتار مشتریان به‌طور قابل‌توجهی تحت تأثیر زمینه‌های مکانی-اقتصادی و روابط همجواری مکانی قرار دارد.

یافته‌های حاصل از مدل‌سازی پیش‌بینی تقاضا (مطابق جداول ۳ تا ۵) نشان می‌دهد که انتخاب مدل مناسب باید فراتر از اتکا به معیارهای کمی صرف بوده و هم‌زمان شواهد تشخیصی مدل، رفتار خطا و قابلیت اتکای احتمالات پیش‌بینی‌شده نیز مدنظر قرار گیرد. در این مطالعه، مدل‌های مبتنی بر درخت مانند Random Forest و XGBoost عملکرد بهتری نسبت به سایر مدل‌ها نشان دادند که این امر ناشی از سازگاری بالای این مدل‌ها با داده‌های مقطعی و روابط غیرخطی است. برخلاف برخی مطالعات پیشین، در این پژوهش مدل ST-GNN عملکرد برتری نسبت به مدل‌های کلاسیک نشان نداد. این نتیجه نشان می‌دهد که کارایی مدل‌های گرافیکی به‌شدت وابسته به وجود داده‌های زمانی و ساختارهای پویای شبکه است و در داده‌های مقطعی، مدل‌های ساده‌تر ممکن است عملکرد بهتری داشته باشند. به‌طور مشابه، عملکرد محدود مدل LSTM نیز بیانگر آن است که استفاده از مدل‌های توالی‌محور بدون وجود داده‌های زمانی مناسب، نمی‌تواند به بهبود معنادار عملکرد منجر شود. علاوه بر این، بهره‌گیری از ابزارهای تبیین‌پذیری مانند SHAP امکان درک بهتر رفتار مدل‌ها و شناسایی عوامل کلیدی مؤثر بر پیش‌بینی‌ها را فراهم می‌سازد و به انتقال نتایج از سطح تحلیل به سطح تصمیم‌سازی مبتنی بر شواهد کمک می‌کند (Lundberg & Lee, 2017). این خروجی‌ها نشان می‌دهند که چگونه می‌توان عوامل کلیدی اثرگذار بر رفتار مشتریان و تقاضای خدمات را شناسایی کرده و از آن‌ها در طراحی سبد خدمات، تخصیص منابع و بازطراحی شبکه شعب استفاده نمود. نتایج این پژوهش با مطالعات اخیر در حوزه GeoAI و هوش تصمیم‌گیری مکانی (Spatial Decision Intelligence) همسو بوده و نشان می‌دهد که ترکیب تحلیل مکانی، یادگیری ماشین و تبیین‌پذیری مدل‌ها می‌تواند زیرساختی مؤثر برای سامانه‌های تصمیم‌یار نسل جدید فراهم آورد [۲۰، ۲۹]. در نهایت، پیوند نتایج پیش‌بینی و خوشه‌بندی مکانی با مسئله مکان‌یابی نشان می‌دهد که تصمیمات مکانی زمانی از منظر علمی و مدیریتی قابل دفاع هستند که بر پایه تقاضای پیش‌بینی‌شده، ساختار همجواری مکانی و عوامل اقتصادی-جمعیت‌شناختی اثرگذار تبیین شوند. از منظر کاربردی، چارچوب پیشنهادی این پژوهش می‌تواند به بانک‌ها کمک کند تا با درک عمیق‌تری از نیازهای محلی، خدمات خود را هدفمندتر و

شخصی‌سازی‌شده‌تر ارائه دهند، هزینه‌های عملیاتی را کاهش دهند و سطح دسترسی و رضایت مشتریان از خدمات بانکی را ارتقا بخشند. با این حال، این پژوهش دارای محدودیت‌هایی نیز می‌باشد. مهم‌ترین محدودیت، مقطعی بودن داده‌ها و عدم دسترسی به داده‌های زمانی پیوسته است که بر عملکرد مدل‌های عمیق و گرافیکی تأثیرگذار بوده است. همچنین، ساختار گراف مورد استفاده در مدل ST-GNN محدود به همسایگی مکانی ساده بوده و تعاملات پیچیده‌تر شبکه‌ای در آن لحاظ نشده است. این پژوهش با هدف توسعه یک چارچوب تصمیم‌یار داده‌محور برای تحلیل تقاضای خدمات بانکی و پشتیبانی از تصمیمات مکانی در شبکه شعب بانک انجام شد. نتایج نشان داد که ادغام تحلیل‌های مکانی، یادگیری ماشین و روش‌های تبیین‌پذیری می‌تواند شناخت دقیق‌تری از الگوهای تقاضای خدمات بانکی فراهم کند. برخلاف مطالعاتی که صرفاً بر مقایسه عملکرد الگوریتم‌ها تمرکز دارند، هدف اصلی این پژوهش ارائه یک چارچوب کاربردی برای شناسایی خوشه‌های تقاضا، استخراج نقاط داغ خدمات بانکی و پشتیبانی از مکان‌یابی شعب، توسعه خدمات دیجیتال و تخصیص بهینه منابع بانکی بوده است.

پیشنهادهای کاربردی و مسیرهای پژوهشی آینده

بر اساس یافته‌های این پژوهش و محدودیت‌ها، چند مسیر اولویت‌دار برای توسعه مطالعات آتی قابل طرح است. نخست، ادغام داده‌های تکمیلی و غنی‌تر، از جمله داده‌های زمانی واقعی (تراکنش‌های تاریخ‌دار در مقیاس روزانه یا ساعتی)، داده‌های اقلیمی، داده‌های رفت‌وآمد شهری و داده‌های رفتاری حاصل از کانال‌های دیجیتال، می‌تواند به بهبود دقت پیش‌بینی تقاضای خدمات بانکی، به‌ویژه در مناطق با نوسانات فصلی یا پویایی بالای تقاضا، کمک شایانی کند. چنین داده‌هایی امکان ساخت توالی‌های معتبر را فراهم کرده و استفاده اصولی‌تر از مدل‌های توالی‌محور نظیر LSTM و ترنسفورمر (Transformer) را ممکن می‌سازد.

دوم، توسعه و به‌کارگیری مدل‌های یادگیری عمیق ترکیبی (Hybrid Deep Learning) و معماری‌های پیشرفته‌تر گرافیکی، نظیر شبکه‌های گرافیکی مبتنی بر توجه (Graph Attention Networks) و مدل‌های مبتنی بر کانولوشن زمانی (Temporal Convolution)، می‌تواند ظرفیت مدل را در تحلیل تعاملات پیچیده مکانی-زمانی و ثبت اثرات غیرخطی و سرریز مکانی افزایش دهد؛ به‌ویژه در شرایطی که داده‌های زمانی پیوسته در دسترس باشد. سوم، گسترش تحلیل عدم قطعیت، کالبراسیون چندکلاسه (فراتر از کلاس غالب) و گزارش شاخص‌های پایداری و استحکام، می‌تواند اعتمادپذیری نتایج را برای کاربردهای عملی و سیاست‌گذاری افزایش دهد. چهارم، پیشنهاد می‌شود خروجی‌های تبیین‌پذیری مدل‌ها، به‌ویژه SHAP، به‌صورت مستقیم به فرآیند سیاست‌گذاری و طراحی سبد خدمات متصل شوند؛ به‌گونه‌ای که قواعد قابل توضیح و مبتنی بر داده برای هر خوشه مکانی استخراج

[8] Nguyen T, Lee J, Park S. Customer behavior prediction in digital banking using deep learning. *Appl Soft Comput*. 2022;114:108–126.

<https://doi.org/10.1016/j.asoc.2022.108731>

[9] Zhou L, Pan S, Wang J, Vasilakos AV. Machine learning on big data: Opportunities and challenges. *Neurocomputing*. 2019;237:350–361.

<https://doi.org/10.1016/j.neucom.2018.07.023>

[10] Yin Y, Zhou D. Behavioral data mining for personalized banking services. *Inf Syst Front*. 2021;23(5):1157–1173.

<https://doi.org/10.1007/s10796-020-10042-x>

[11] Ghosh P. Machine learning in banking: Applications and challenges. *J Financ Anal*. 2020;18(3):45–62.

<https://doi.org/10.3920/S1544-709820-0001-5>

[12] Anselin L. Local Indicators of Spatial Association—LISA. *Geogr Anal*. 1995;27(2):93–115.

<https://doi.org/10.1080/10334674.1995.11102731>

[13] Li X, Chen Y, Zhao L. Spatial analytics for banking branch optimization. *Expert Syst Appl*. 2020;145:113130.

<https://doi.org/10.1016/j.eswa.2020.113130>

[14] Hernandez J, Li F, Wang S. Spatial data mining in financial services. *Int J Geogr Inf Sci*. 2019;33(4):722–738.

<https://doi.org/10.1080/15387216.2018.1531234>

[15] Rahman K, Shukla S. Smart location-allocation modeling in service industries using machine learning. *Decis Support Syst*. 2022;160:113756. <https://doi.org/10.1016/j.dss.2022.113756>

[16] Chen X, Wang Y, Li Z. Integrating spatial analytics and machine learning for location-based decision support. *Int J Geogr Inf Sci*. 2020;34(11):2191–2212.

<https://doi.org/10.1080/13658816.2020.1762434>

[17] Wanke P, Barros CP. Location efficiency of bank branches: A network DEA approach. *Omega*. 2014;46:1–14.

<https://doi.org/10.1016/j.omega.2013.07.004>

[18] Chen T, Guestrin C. XGBoost: A scalable tree boosting system. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York: ACM; 2016. p. 785–794.

<https://doi.org/10.1145/2939672.2939785>

[19] Bzdok D, Altman N, Krzywinski M. Statistics versus machine learning. *Nat Methods*. 2021;18:233–234.

<https://doi.org/10.1038/s41592-020-01018-x>

[20] Li W, Hsu CY. GeoAI for intelligent spatial analysis and urban applications. *ISPRS Int J Geo-Inf*. 2022;11(4):248.

<https://doi.org/10.3390/ijgi11040248>

[21] Reis EA, Pacheco A, Santos M. Spatial machine learning: concepts, applications and opportunities. *Environ Model Softw*. 2021; 145:105206.

<https://doi.org/10.1016/j.envsoft.2021.105206>

و در تصمیمات مدیریتی مورد استفاده قرار گیرد. در نهایت، آزمون چارچوب پیشنهادی در مقیاس‌های جغرافیایی متفاوت، در بانک‌ها و مناطق مختلف، و حتی در سایر صنایع خدماتی نظیر بیمه، حمل‌ونقل و خرده‌فروشی، می‌تواند امکان ارزیابی قابلیت تعمیم، پایداری و اثربخشی این رویکرد داده‌محور را فراهم آورد.

مشارکت نویسندگان

علی رضائی این مقاله را از رساله دکتری خود، تالیف نمود. محمد قاسمی، نقش ویرایش و نظارت بر مقاله و محمد تفرشی ویرایش و مشاوره را برعهده داشتند.

تشکر و قدردانی

این مقاله حاصل خروجی‌های تحقیقاتی از رساله دکتری علی رضائی با راهنمایی و نظارت محمد قاسمی و مشاوره سیدمحمد تفرشی می‌باشد.

تعارض منافع

«هیچ‌گونه تعارض منافع توسط نویسندگان بیان نشده است.»

منابع و مآخذ

[1] Chen H, Chiang RHL, Storey VC. Business intelligence and analytics: from big data to big impact. *MIS Q*. 2012;36(4):1165–1188. doi:10.2307/41703503.

[2] McKinney E, Yoos CJ, Snead KC. The need for speed: An analysis of enterprise data analytics and decision-making. *Decis Support Syst*. 2017;99:1–14. <https://doi.org/10.1016/j.dss.2017.03.001>

[3] Kumar R, Lim C. AI-based dynamic banking services: A predictive approach. *J Intell Syst*. 2023;29(2):101–120. <https://doi.org/10.1007/s10845-022-09463-w>

[4] Bhatore S, Mohan L, Rathee G. Artificial intelligence and machine learning techniques for smart banking systems: a review. *Int J Inf Manage Data Insights*. 2020;1(1):100001. <https://doi.org/10.1016/j.ijime.2020.100001>

[5] Cao L. AI in finance: challenges, techniques and opportunities. *ACM Comput Surv*. 2022;55(3):1–38. <https://doi.org/10.1145/3503253>

[6] Dwivedi YK, Hughes L, Ismagilova E, Aarts G, Coombs C, Crick T, et al. Artificial Intelligence (AI): multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research. *Int J Inf Manage*. 2021;57:101994. <https://doi.org/10.1016/j.ijinfomgt.2020.102169>

[7] Bose I, Kumar V, Gupta V. Artificial intelligence in financial services: a systematic review and future research agenda. *Int J Inf Manage Data Insights*. 2023;3(2):100192. <https://doi.org/10.1016/j.ijime.2023.100192>

[34] Minderer M, Djolonga J, Romijnders R, Hubis F, Zhai X, Houlsby N, et al. Revisiting the calibration of modern neural networks. *Adv Neural Inf Process Syst*. 2021;34:15682-15694.

[35] Davis J, Goadrich M. The relationship between precision-recall and ROC curves. In: *Proceedings of the 23rd International Conference on Machine Learning*. New York: ACM; 2006. p. 233-240. <https://dl.acm.org/doi/10.5555/1143846.1143873>

[36] Saito T, Rehmsmeier M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS One*. 2015;10(3): e0118432. <https://doi.org/10.1371/journal.pone.0118432>

[37] Niculescu-Mizil A, Caruana R. Predicting good probabilities with supervised learning. In: *Proceedings of the 22nd International Conference on Machine Learning*. New York: ACM; 2005. p. 625-632. <https://dl.acm.org/doi/10.5555/1096875.1096952>

[38] Murray AT. Advances in location modeling: GIS linkages and contributions. *J Geogr Syst*. 2010;12(3):335-354. <https://doi.org/10.1002/jgs.731>

[22] Janowicz K, Gao S, McKenzie G, Hu Y, Bhaduri B. GeoAI: Spatially explicit artificial intelligence techniques for geographic knowledge discovery and beyond. *Int J Geogr Inf Sci*. 2020;34(4):625-636. <https://doi.org/10.1080/13658816.2019.1723832>

[23] Zhou J, Cui G, Hu S, Zhang Z, Yang C, Liu Z, et al. Graph neural networks: a review of methods and applications. *AI Open*. 2020;1:57-81. <https://doi.org/10.1016/j.aiopen.2021.05.001>

[24] Wu Z, Pan S, Chen F, Long G. Graph neural networks in spatiotemporal forecasting: recent advances and future directions. *Mach Intell Res*. 2023;20(4):451-474. <https://doi.org/10.1007/s11468-023-01844-w>

[25] Chen X, Wang Y, Li Z. Integrating spatial analytics and machine learning for location-based decision support. *Int J Geogr Inf Sci*. 2020;34(11):2191-2212. <https://doi.org/10.1080/13658816.2020.1733423>

[26] Shwartz-Ziv R, Armon A. Tabular data: deep learning is not all you need. *Inf Fusion*. 2022;81:84-90. <https://doi.org/10.1016/j.inffus.2021.11.006>

[27] Kipf TN, Welling M. Semi-supervised classification with graph convolutional networks. In: *International Conference on Learning Representations (ICLR)*; 2017. <https://doi.org/10.48550/arXiv.1609.02907>

[28] Jiang D, Luo C. Graph neural network for spatio-temporal forecasting: a survey. *Artif Intell Rev*. 2022;55:6051-6084. <https://doi.org/10.1007/s10462-022-10168-x>

[29] Delamater PL, Messina JP, Shortridge AM, Grady SC. Measuring geographic access to health care: raster and network-based approaches. *Health Place*. 2022;73:102745. <https://doi.org/10.1016/j.healthplace.2022.102745>

[30] Arrieta AB, Díaz-Rodríguez N, Del Ser J, Bennetot A, Tabik S, Barbado A, et al. Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf Fusion*. 2020;58:82-115. <https://doi.org/10.1016/j.inffus.2019.12.012>

[31] Roscher R, Bohn B, Duarte MF, Garcke J. Explainable machine learning for scientific insights and discoveries. *IEEE Access*. 2020;8:42200-42216. <https://doi.org/10.1109/ACCESS.2020.2976199>

[32] Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: *Advances in Neural Information Processing Systems (NeurIPS)*; 2017. p. 4765-4774. <https://doi.org/10.48550/arXiv.1705.07877>

[33] Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, et al. From local explanations to global understanding with explainable AI for trees. *Nat Mach Intell*. 2020;2(1):56-67. <https://doi.org/10.1038/s42256-019-0138-9>

معرفی نویسندگان

AUTHOR(S) BIOSKETCHES



علی رضایی دانشجوی دکتری مدیریت بازرگانی گرایش بازاریابی دانشگاه تربت حیدریه می‌باشد. ایشان مدرک کارشناسی ارشد مدیریت بازرگانی گرایش بازرگانی داخلی خود را در دانشگاه آزاد واحد اراک در سال ۱۳۹۵ و مدرک کارشناسی مدیریت بازرگانی خود را در سال

۱۳۹۳ در دانشگاه پیام نور اراک اخذ نموده است. از سال ۱۴۰۲ به عنوان دستیار آموزشی دانشگاه آزاد واحد کمیجان و تفرش و از سال ۱۴۰۴ به عنوان دستیار آموزشی دانشگاه آزاد واحد اراک مشغول به فعالیت می‌باشد. کتاب بررسی استقرار عوامل مدیریت دانش در بانک در سال ۱۴۰۰ به قلم ایشان به چاپ رسیده است.

Rezaei A. PhD Student in Management, Department of Management, ToH.C., Islamic Azad University, Torbat Heydarieh, Iran.

✉ eilia.rezaei@iau.ac.ir



محمد قاسمی نامقی استادیار دانشگاه پیام نور واحد مشهدگروه مدیریت بازرگانی. ایشان مدرک دکترا و کارشناسی ارشد خود را به ترتیب در رشته مدیریت بازرگانی و رشته مدیریت اطلاعات IT سال ۱۳۸۷ و ۱۳۸۲ در دانشگاه علامه طباطبائی دریافت نموده اند و

مدرک کارشناسی خود را سال ۱۳۷۹ در رشته مدیریت بازرگانی در دانشگاه ولی عصر (عج) رفسنجان اخذ نموده اند. تاکنون ایشان ۲۰ مقاله علمی در زمینه های مدیریت و حسابداری در مجلات مختلف به



گرایش تصمیم‌گیری و خط‌مشی‌گذاری عمومی از دانشگاه آزاد کرمان سال ۱۳۹۸ اخذ نموده است. تاکنون مقالات متعدد مدیریتی و حسابداری را به چاپ رسانده اند که اخیراً مقالات بررسی چشم‌انداز کاری با بررسی رضایت مشتریان در شبکه نسل سوم

اینترنت موبایل چاپ شده در مجله پژوهش‌های علوم مدیریت به چاپ رسانده است.

Tafreshi S.M. Assistant Professor, Department of Management, ToH. C., Islamic Azad University, Torbat Heydarieh, Iran.

✉ tafreshiazad1@iau.ac.ir

چاپ رسانده اند که آخرین مقاله ایشان کاربرد چارچوب مفهومی بازاریابی اجتماعی در رفتارهای سلامت در فصلنامه نوید نو (دانشگاه علوم پزشکی مشهد) می‌باشد.

Ghasemi M. Assistant Professor of Management, Department of Economics, Management and Accounting, Payame Noor University, Tehran, Iran.

✉ ghasemi.mo@pnu.ac.ir

سید محمد تفرشی استادیار تصمیم‌گیری و خط‌مشی‌گذاری عمومی دانشگاه آزاد واحد تربت حیدریه. ایشان مدرک لیسانس حسابداری از دانشگاه صنعت نفت تهران در سال ۱۳۷۵ و کارشناس ارشد مدیریت بازرگانی از دانشگاه فردوسی مشهد و دکترای مدیریت دولتی

Citation (Vancouver): Rezaei A, Ghasemi M, Tafreshi S.M. [Development of Data-Driven Framework Based on Spatial Analysis and Graph Neural Networks for Service Demand Forecasting and Optimization of Banking Branch Networks]. *J. RS. GEOINF. RES.* 2026; 4(1): 107-134

 <https://doi.org/10.22061/jrsgr.2026.113044.1124>



COPYRIGHTS

© 2026 The Author(s). This is an open-access article distributed under the terms and conditions of the Creative Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) (<https://creativecommons.org/licenses/by-nc/4.0/>)