Localinsights: Areawise Business Analysis

Rashmi Kumari Raj Kapoor Choudhary[1], Sakshi Dattatray Jadhav[2], Shubham Deepak Jamdar[3], Aryan Sandeep Bhekare[4], Prof. Vilas Jadhav[5]

[1,2,3,4] *Computer Engineering Mahatma Gandhi Mission's College of Enginnering and Technology Navi Mumbai, India*

[5] *Dept. Computer Engineering Mahatma Gandhi Mission's College of Engineering and Technology Navi Mumbai, India*

***Abstract*—One of the most important and difficult choices for local business owners and entrepreneurs is choosing a suitable location. These decisions are often based on insufficient information,inadequate market knowledge, or personal intuition, which frequently leads to inadequate planning and financial setbacks. Despite the abundance of location-based and business-related data available, it is usually distributed across multiple platforms and presented in complex formats, making it difficult for non-technical users to comprehend and use efficiently. To overcome these limitations, this paper introduces Local Insight AI, a web-based decision support system designed to enhance and expedite the business location selection within the state of Maharastra. By mixing descriptive market scoring techniques with predictive machine learning models, the platforms create clear visual reports and comparative analyses across different regions.**

***Index Terms*—Business Decision Support System, Location Intelligence, Geospatial Analysis, Artificial Intelligence, Machine Learning Model, LightGBM.**

I. INTRODUCTION

1.1 Context and Motivation

In quickly growing urban areas and support mercenary environments, selecting the success business growth. can a result, role in determining the success or failure of the abominable A well-chosen location selection importantly increase customer footfall, improve visibility, and small tall term business location On the most hand, an abominable location selection led to low customer footfall, operable difficulties, and small losses.

With the potential of geographical technologies, rhetorical amounts of geographical, competition, in infrastructure related data are now available. This includes population density, purchasing behaviour, in intensity, accessibility, competition business details. Most mingy entrepreneurs and used business owners lack the need required to account and combine this data driven.

1.2 Research Gap and Challenges

A thorough review of existing literature shows that significant research has been done in business location analytics, artificial intelligence, learning-to-rank models, and decision support systems. However, despite these improvements, small and medium-sized enterprises (SMEs) still have limited use of such systems [1], [6]. Many studies focus on complex AI-based ranking models and intelligent systems for large companies [2], [3]. Several research works highlight how business location affects sales performance and profitability [4], [10], [11]. However, in practice, selecting a location is mostly based on intuition. Many small business owners rely on personal experience, manual field surveys, or local opinions instead of systematic data analysis [13], [17].

Even though large amounts of geographical data, demographic data, online reviews, and urban big data are available today [9], [13]. Extracting useful insights from such disjointed data requires technical skills, making effective analysis challenging for non-technical users [8], [20]. Existing tools mainly focus on visualization techniques, including maps, dashboards, and graphical-reports [11]. Another major limitation is that many AI-driven or multi-criteria decision-making models require theoritical knowledge of algorithm like MCDM, ranking techniques, or machine learning methods [1],[5],[16].

1.3 Research Objectives and Key Contributions

The primary objectives and contributions of this research work are described in the following:

- To create and build Local Insight AI, an AI-powered decision support system for business location identification.
- To incorporate various location-related and business data into a single analytical system.
- To offer location comparison and recommendation services based on specific business requirements.

II. LITERATURE REVIEW

1.AI and Big Data for Entrepreneurial and Strategic Decision Making

In[1], the authors use a Multi-Criteria Decision-Making framework and structured evaluation techniques to choose suitable Explainable AI methods. The study in [2] looks at Learning-to-Rank models that use LightGBM. It shows how gradient boosting-based ranking improves accuracy in priritization for large-scale intelligent systems.Similarly, [3] examines Learning-to-Rank algorithms like SVMRank and pairwise ranking methods.In contrast, [4] discusses the use of AI-driven predictive analytics and big data modeling to aid enterprenurial decision-making in the digital economy. While these studies clearly show the strengths of MCDM frameworks, ranking algorithms, and predictive analytics, they mainly focus on the technical performance of models.

2.AI-Based Decision Support Systems and Business Analytics

Hjelle [5] discusses how organizations make decisions using analytics-driven frameworks and decision-support systems (DSS). She highlights the use of data modeling and structured analytical tools to improve managerial accuracy. Al-Surmi [6] suggests an AI-based hybrid decision model that combines different computing strategies to improve performance and lower uncertainty in business evaluation.

Badmus et al. [7] examine AI-driven business analytics with machine learning algorithms, predictive modeling, and automated decision frameworks. Similarly, Islam [8] introduces a Real-Time AI-Based Informational Decision Support system that relies on smart data processing and automated analysis modules to aid dynamic decision-making.

3. Location Intelligence and Business Site Selection

Arshad [9] highlights how Big Data analytics and AI-based data processing frameworks help us understand online market behavior and performance patterns. Liang et al. [10] introduce the Dynamic Huff Model, which uses location big data and spatial interaction modeling to estimate business attractiveness and customer flow in different geographic areas.

Balumbach et al. [11] propose QUIS, an AI-supported framework that brings together geospatial analysis, structured data processing, and smart ranking techniques for automated business site assessment. D'Silva et al. [12] contribute further by applying spatio-temporal data analytics and temporal activity predictions models to study demand changes for new venues over time.

4. Explainable, Trustworthy, and Human-Centred AI

Xiang et al. [13] use urban big data analytics, GIS based spatial modeling, and structured data integration frameworks to support business location decisions. Alam et al. [16] demonstrate how spatial clustering algorithms, regression-based spatial modeling and geographic feature extraction techniques improve reliablity in location-based decisions. Liang et al. [10] (Dynamic Huff Model) use probabilistic spatial interaction theory and location big data calibration to improve business attractiveness modeling.

5. Web-Based and Hybrid Intelligence Decision Platform

Baumbach et al. [17] introduce QUIS, an AI-enabled decision framework that combines web-based geospatial analytics, structured data processing, and intelligent ranking systems to evaluate business sites. D'Silva et al. [18] use spation-temporal activity modeling and predictive time series analysis to forecaste venue demand patterns.

Xiang et al.[19] use urban big data analytics,GIS based spatial intelligence, and demographic data fusion techniques to improve location-aware decisions systems. The study emphasizes the benefits of integrating deiverse web-sourced datasets into a single analytical framework for business planning.

Machireddy et al. [20] further stress the importance of machine learning algorithms, regression modeling, classification techniques, and business intelligence platforms.

Table 1: Comparative analysis of existing Research Works

Ref.	Method Used	Tools & Technologies	Strengths	Limitations
[2]	Learning-to-Rank using Gradient Boosting	LightGBM, ML ranking algorithms	High scalability, strong ranking accuracy	Complex implementation, less explainability
[10]	AI and Big Data Analytics for decision support	AI models, Big Data frameworks	Supports entrepreneurial decisions with data-driven insights	Focused on digital economy, not location-specific
[12]	Hybrid AI decision-making strategies	AI-based optimization and analytics tools	Combines multiple AI strategies to improve performance	Limited practical implementation details
[16]	Dynamic Huff Model	Location Big Data, spatial modeling	Strong for business location analysis	Traditional model, lacks AI-based ranking
[17]	Intelligent Site Selection System	Knowledge-based systems, spatial analytics	Structured site selection approach	Limited scalability, less real-time capability

The research shows that AI, big data analytics, and spatial modeling significantly improve decision accuracy and strategic planning.

III. PROPOSED METHOD

3.1 Architecture of the System

As shown in Figure 1, The LoacalInsight AI architecture uses a layered design that combines user interaction backend processing, data management, and AI-based analysis into one system.

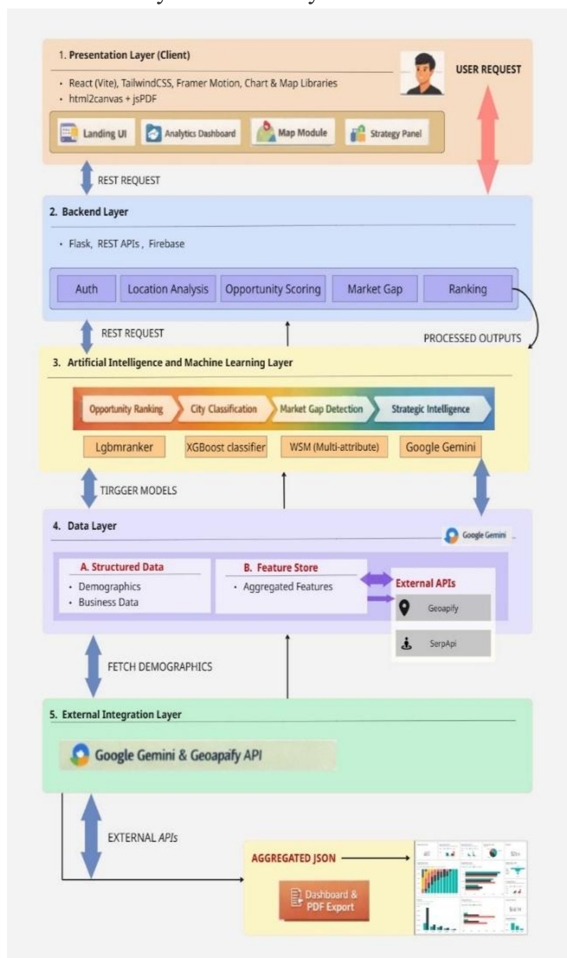


Fig. 1: Local Insight AI System Architecture Diagram

The web-based user interface allows users to enter business and location queries using a simple, intuitive design. The data layer stores and cleans locations, demographic and business data. The machine learning layer assesses location suitability based on factors like demand, population and compitition. The presentation layer shows ranking and explanation in clear visual format.

3.2. Data Collection Framework

Data gathered from trustworthy public sources to support model training and evaluation. This includes demographic, economic, business and geographic datasets to evaluate different business environments.

3.2.1. Data Acquisition

Demographic and locality -level data were obtained from platform offer organized insights into population distribution, locality characteristics, and area based attributes. Business density and conts of establishment by category, extracted through structured web scraping of publicly available listing on Google maps and other business directories. Due to the heterogenous nature of the collected datasets. A systematic preprocessing stage was implemented to ensure data quality.

3.2.2. Data Preprocessing

Due to the heterogenous nature of the collected datsets, a systematic preprocessing stage was implemented to ensure data quality and consistency.

The preprocessing includeed:

- Dealing with missing values using suitable imputation techniques.
- Normalizing numerical attributes to remove scale bias.
- Encoding categorical variables for model compatibilty.

3.2.3. Target Variable Construction

Since publicly available datasets do not show direct financial indicators like revenue or profit at the local level , a proxy measure is created to estimate business success. It is influenced by client satisfaction, customer angagement and perceived service quality. Studies also shows online ratings and reviews significantly impact consumer behaviour and offer insight into market perception [8,9]. Based on this, the success ration feature is developed:

$$\text{Success Ratio}=\frac{N_{high\_rating}}{N_{total}}$$

Were,

- N_{high_rating} represents the number of establishments in a specific category within an area that have a rating above 4.0 and receive more than 50 customer review.
- N_{total} represents the total number of establishments in that category within the same area.

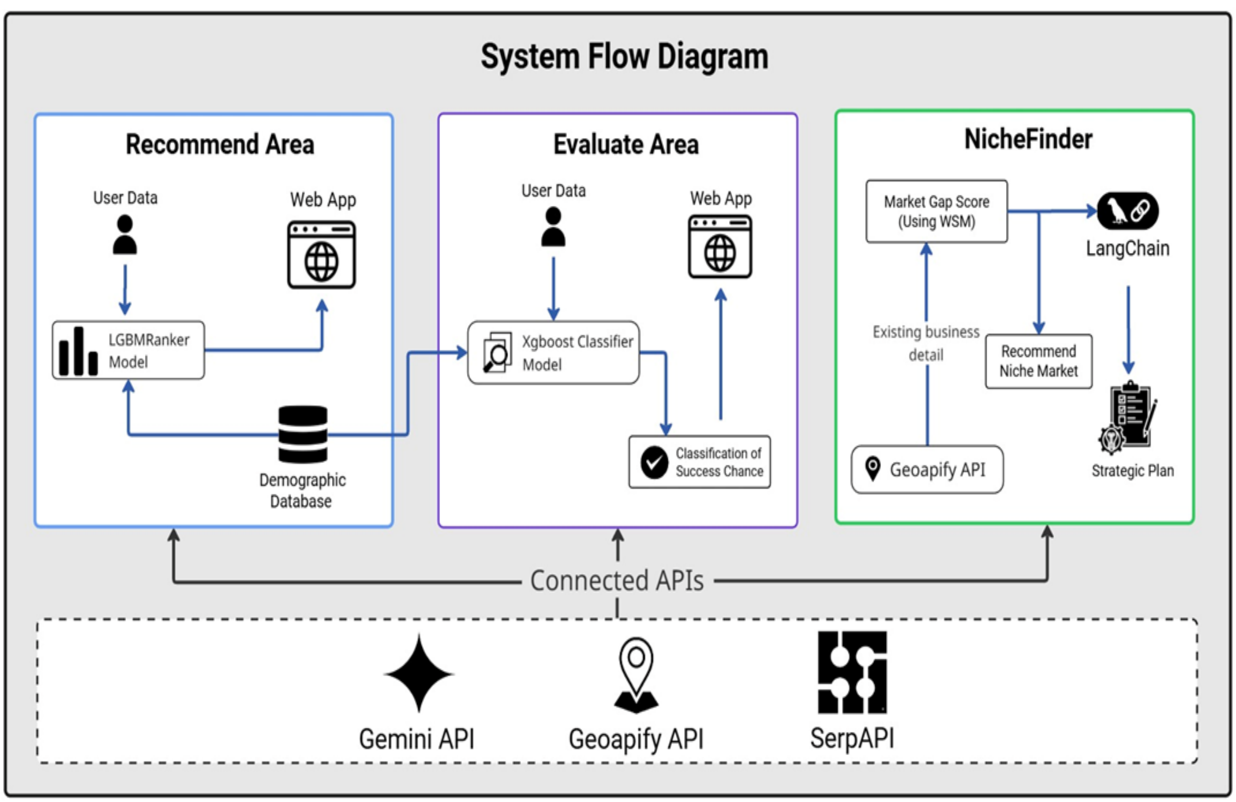


Fig. 2: Proposed System Flow Digram of the Local Insight AI platform

3.3. Implementation Details

The proposed system combines three analytical components into a single framework that supports business decisions:(1) Learning-to-rank for recommending areas, (2) Supervised classifications for evaluating city suitability and (3) multi criteria market gap analysis for identifying niches.

3.3.1 Recommend Area: Contextual Learning-to-Rank Framework

As shown in Figure 2, the Recommend Area module approaches location recommendations as a grouped learning-to-rank problem using LGB MRanker.The model learns the relative ranking of geographic areas within a specific business category and target customer segment.

The ranking goal relies on the SuccessRatio metric, which shows the share of highly rated establishments within a category and area. LGBMRanker is chosen for its computational efficiency, scalability, and solid performance with structured tabular data.[3,6] The system provides the top three ranked areas to ensure clarity and efficient decision-making.

3.3.2 Evaluate City: Suitability Classification Framework

The Evaluate Area module conducts supervised classification to determine how suitable a selected city is for a specific business category. An XGBoost Classifier, is trained using demographic details, business density, and performance-related features. The classifier produces posterior probabilities, which are then categorized into three levels of suitability:High, Medium, Low.

3.3.3 NicheFinder: Market Gap Optimization Framework

This module finds undeserved business opportunities in a chosen city using a multi-criteria decision-making (MCDM) framework. MCDM methods are frequently used in complex situation where different criteria need to be evaluated at the same time.[5]Niche discovery is treated as a multiattribute optimization problem that balances demand indicators, competitive desity and demographic alignment.

Niche potential is measured using a weighted sum model, also known as the Simple Addictive Weighting method, which is one of the oldest and most recognized MCDM techniques.[7] The composit score is defined as:

Score = $\sum_{i=1}^{n} \omega_i f_i$

Were,

- f_i represent normalized market indicators.
- ω_i denotes their respective importance weights.

Competitive density, which is the number of existing businesses within a category compared to the overall size of the city , is included to reflect potential market saturation or unmet demand.

IV. EVALUATION

4.1 Recommendation Model Evaluation

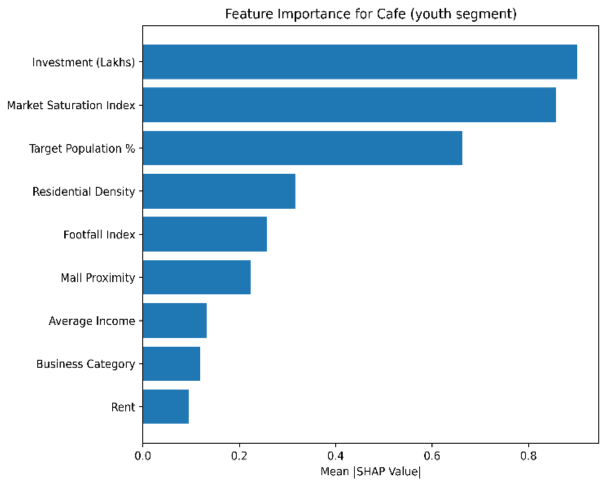
To improve the recommendation framework, look at three approches: a rule-based Weighted Sum Model, a Random Forest Regressor trained on the proxy Success Ratio target and the suggested grouped LGBRanker.

As shown in Table 2, both WSM and Random Forest use global feature weighting and do not show difference across business categories. In contrast , the grouped LGBRanker captures category-specific ranking patterns and achieves an NDCG@3 score of 0.7304.

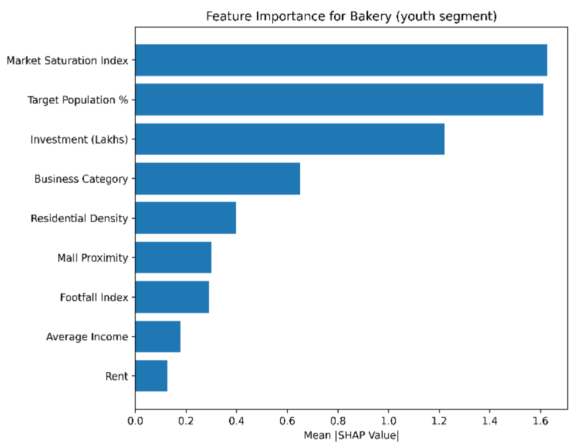
Table 2: Comparison of Recommendation Models

Model	Approach Type	Context-Aware	Key Limitation
WSM	Rule-based heuristic	No	Fixed weights, no learning
Random Forest Regressor	Regression (predict-then-rank)	No	Global feature importance across groups
LGBM Ranker	Grouped Learning-to-Rank	Yes	Requires sufficient data within each group

Comparison shows that Traditional models lack context awareness and don't learn adaptively. LGBM Ranker is context-aware, but it requires enough grouped data.



(a) Cafe Category and Youth Segment



(*b*)*Bakery Category and Youth Segment*

Fig. 3: Feature Importance Variation Across Business Categories

Feature importance differed among business categories and customer segments. The category-wise importance distributions are shown in Figure 3.

4.2 Suitability Prediction Model Evaluation

To evalute the predictive performance of the success classification task, study of three machine learning classifiers is considered.

Table 3: Comparative Analysis of Ensemble Models

Model	Algorithm Type	Strength
XGBoost Classifier	Gradient Boosting	Handles complex nonlinear patterns
Random Forest Classifier	Bagging Ensemble	Robust to overfitting
LightGBM Classifier	Gradient Boosting	Fast and scalable

All three ensemble models capture complex patterns. LightGBM stands out for its speed and scalability. Random Forest provides robustness and XGBBoost delivers strong nonlinear modeling.

Table 4: Comparison of Classification Models

Model	Accuracy	Precision	Recall	F1-Score
XGBoost	94.57	92.70	90.75	91.68
Random Forest	71.09	65.65	58.93	61.14
Light GBM	94.35	91.71	92.05	91.82

As shown in Table 4, ensemble -based boosting models, XGBBoost and LightGBM, showed better predictive.The ability than Random Forest.

V. RESULT AND DISCUSSIONS

The performance shows that Local Insight AI outperforms the baseline method in accuracy and speed. This improvement happens because the system analyzes location, population and market data together rather than separately.

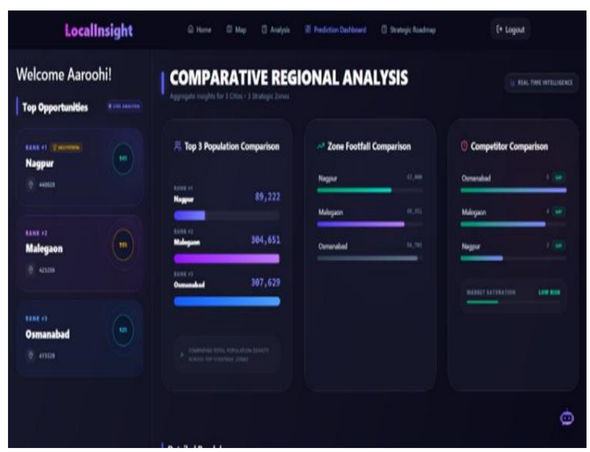


Fig. 4: Comparative Regional Analysis Dashboard

The dashboard provides a straightforward comparison of regions based on population, foot traffic and competition as shown in Fig 4.

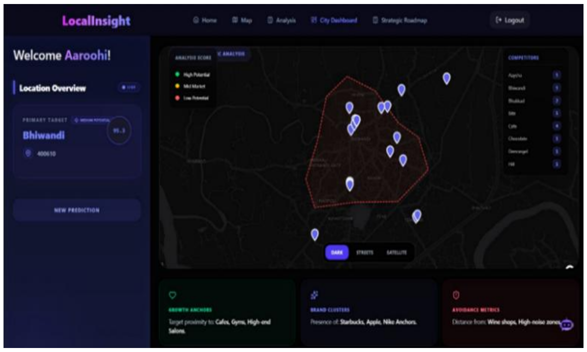


Fig. 5: City Intelligence Dashboard

In Fig. 5, The dashboard offers a detailed analysis at the city level. This tool helps users spot high-potential areas with clear visuals and location-based data.

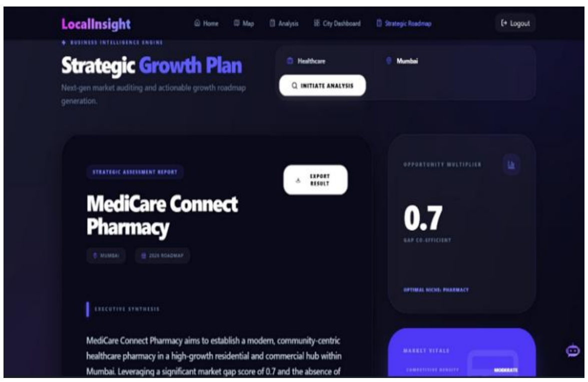


Fig. 6: Local Insight AI- Strategic Growth Plan Dashboard

The dashboard creates a growth plan based on analyzing market gaps and opportunities. It offers practicle insights to help business plan their expansion clearly and confidently.

VI. CONCLUSION AND FUTURE SCOPE

In this research, the LocalInsight AI system was creates as a helpful decisionsupport tool for business owners to make better location choices using data and artificial intilligence. The tool combines location data, machine learning analysis and strightforward web interface to improve experience of making location decision.

Future work can focus on integrating realtime data streams to allow for more dynamic and current location analysis. The system can be improved with deep learning models to further boost prediction accuracy and recommendation quality.

REFERENCES

[1] "A Multi-Criteria Decision-Making Approach for the Selection of Explainable AI Methods," Machine Learning and Knowledge Extraction, vol. 7, no. 4, Art. no. 158, 2025.

[2] "Leveraging Learning-to-Rank Models Using LightGBM for Large-Scale Intelligent Systems," in Proceedings of the ACM Conference, 2024, doi: 10.1145/3687151.3687156.

[3] M. Shuaib, Learning-to-Rank Models for Intelligent Decision Systems, Master's Thesis, BHT Berlin, 2023.

[4] "The Effect of Business Location on Sales Result in the Traditional Market," ResearchGate, 2021.

[5] "A Survey on Multi-Criteria Decision-Making Methods and Its Applications," Academia.edu, 2017.

[6] M. Shuaib, Learning-to-Rank Models for Intelligent Decision Systems, Master's Thesis, BHT Berlin, 2023.

[7] "Leveraging Learning-to-Rank Models Using LightGBM for Large-Scale Intelligent Systems," ACM Digital Library, 2024, doi: 10.1145/3687151.3687156.

[8] "Impact of Consumer-Generated Online Reviews and Ratings on Purchase Behavior and Sales Performance," International Journal of Research and Scientific Innovation (IJRSI), 2024.

[9] "The Influence of Online Reviews on Consumer Behaviors and Purchasing Decisions: A Narrative Review," ResearchGate, 2024.

[10] Y. Cao, "Using AI and big data analytics to support entrepreneurial decisions in the digital economy," Scientific Reports, vol. 15, Art. no. 36933, 2025.

[11] S. Hjelle, "Organizational decision making and analytics," Decision Support Systems, 2024.

[12] A. Al-Surmi, "AI based decision making: combining strategies to improve performance," International Journal of Technology Management, 2022.

[13] O. Badmus, S. A. Rajput, J. B. Arogundade, and M. Williams, "AI-driven business analytics and decision making," World Journal of Advanced Research and Reviews, vol. 24, no. 1, pp. 616–633, 2024.

[14] A. Islam, "Real-Time AI-Based Informational Decision-Making Support System," Applied Sciences, vol. 11, no. 13, Art. no. 6237, 2021.
[15] M. Arshad, "Big data analytics and AI as success factors for online business," PMC, 2025.
[16] Y. Liang, S. Gao, Y. Cai, N. Zhang Foutz, and L. Wu, "Calibrating the dynamic Huff model for business analysis using location big data," arXiv preprint, 2020.
[17] S. Baumbach, F. Sachs, S. Ahmed, and A. Dengel, "QuIS: The Question of Intelligent Site Selection," arXiv preprint, 2020.
[18] K. D'Silva, A. Noulas, M. Musolesi, C. Mascolo, and M. Sklar, "Predicting the temporal activity patterns of new venues," arXiv preprint, 2018.
[19] Y. Xiang, D. Chang, and X. Feng, "Leveraging urban big data for informed business location decisions," arXiv preprint, 2023.
[20] R. Machireddy, R. Rachakatla, and P. Ravichandran, Advanced Business Analytics with AI, Science Academia Press, 2023.