Framework For Leveraging A.I. for Strategic Growth in Small-Scale Enterprises

Susmit Acharya[1], Soma Chakraborty (advisor)[2]
[1,2]*Department of Computer Science, Auxilium Convent School, Barasat, Kolkata, India*

***Abstract* — This research explores low-cost, AI-based business intelligence (AI-BI) platforms customized for Small Scale Enterprises around Kolkata, bridging infrastructure, budget, and digital preparedness constraints. With open-source machine learning software, an AI-BI dashboard prototype using Python libraries (Scikit-learn, XGBoost, FastAPI) was created and tested with synthetic data mimicking retail and manufacturing activities. Central functionalities were demand forecasting, churn prediction, and customer segmentation. Simulation evaluation revealed improved mean forecast accuracy by 35% and clearer segmentability of customers. Although the use of real business data for the models was not possible due to anonymity concerns considering the open-source nature of the research, the project contributes a reproducible, open-source platform design and evaluation framework. Results prove AI-BI adoption at scale as technically feasible in under-resourced contexts of SMEs and provide foundations for future field verifications.**

***Index Terms* — Robotics and Intelligent Machines; Machine Learning; Ai-Based Business Intelligence; Small-Scale Enterprises; Synthetic Data; Framework; Open-Source**

1. INTRODUCTION

1.1 Context of the Research

In India, Small Scale Enterprises make up about roughly 30% of the GDP [1] and employ over 120 million people [2]. In the state of West Bengal, there is a large number of Small-Scale Enterprises in textile, leather, retail, and service sectors. However, according to data figures, only about 15% of Indian SMEs have integrated Digital or AI-based tools into their operations [3]. This can be attributed to cost-sensitivity of existing technologies, limited awareness about use-case, complicated implementation methods, and unsure Return on Investment [4]. The demand for cost-effective and easy-to-implement digital-solutions is rising among these enterprises, but implementation still remains a problem due to above mentioned reasons.

1.2 The Problem

Despite the potential of Artificial Intelligence to streamline business processes, small scale enterprises in Kolkata struggle to access and implement low-cost, easy-to-implement AI tools to automate tasks. The tools and frameworks that currently exist (like TOE framework [5] or TAM framework [6]) are often unaffordable, overly complex, or require other technologies available together to give accurate results [7]. These frameworks lack adequate local feasibility study especially in cost-sensitive and low-experience enterprises [8]. This leads to reliance on instinct-based or manual decision-making.

1.3 Background and Related Studies

The existing work researching AI adoption among enterprises has investigated several conceptual frameworks. Common Artificial Intelligence Maturity Model (AIMM [9]) proposes a five-staged framework— Awareness, Experimentation, Operational, Optimization, Transformation —to frame AI adoption as a continuum [10]. Combining structural preparedness factors (Technology–Organization–Environment [5]), behavioral intention models (Technology Acceptance Model [6]), and operational stage (AIMM [9]) would encapsulate adoption dynamics in a more complete framework.

Various researchers urge caution in excessive reliance on imported models, emphasizing localized solution approaches so that they align with local conditions, such as infrastructure limitations, educational level, and industry-specific workflow. A recent study by Serena Proietti and Roberto Magnani in 2025 [11] explicitly states that “SMEs face considerable obstacles, largely due to the smaller amounts of data they handle and the generally lower quality of the data they have available” which is a huge problem when it comes to globalized frameworks. Another study in

2024 [12] provides a comparative analysis of regional barriers (financial, skill, regulatory) across continents and highlights how high-cost, low-experience" barrier is a common, yet under-studied, problem in emerging economies like India.

Measuring AI's impact has also been erratic, as metrics used differ substantially between studies. Most sophisticated analyses use quantitative KPIs [13], such as:

KPI	Description	Typical Improvement
Forecast Accuracy	Mean Absolute Percentage Error below 20%	+30–40% predictive value
Downtime Reduction	Reduction in unplanned system or process downtime	Up to 40%
Customer Satisfaction	Improved NPS and customer retention metrics	10–20% uplift
Revenue Uplift	Gains through AI-led targeting and dynamic pricing	5–10% growth

Despite stated advances, there is a lack of study for a localized approach for an AI-BI framework for small-scale enterprises.

1.4 Identified Research Gap

In the recent studies in this field, there is a lack of specific focus on AI adoption in Eastern India. Additionally, studies regarding feasibility of existing frameworks for cost-sensitive small-scale enterprises are inadequate. Also, a comprehensive study on localized AI tools in local small businesses is mostly unexplored.

1.5 Objectives of the Research

To identify practical, cost-effective AI strategies that address common pain points in workflows of these small-scale enterprises.

To design and prototype an open-source AI business intelligence dashboard that is tailored specifically to the needs of these and similar enterprises.

To evaluate business performance improvements in quickly-measurable areas like sales forecast accuracy, customer-segmentation to prioritize the customers bringing in the highest revenue, and customer satisfaction.

1.6 Scope and Limitations

While comprehensive, this research data was limited to small-scale enterprises with fewer than 250 employees and only five sectors were selected: manufacturing, retail, healthcare, logistics, and finance. All the enterprises were in the greater Kolkata region which might affect the accuracy of the data presented for other regions. Another limitation of this study is the synthetic data used for the results. However, the data was created after careful analysis of 120+ small-scale enterprises in the greater Kolkata region and thus tries its best to mimic real business data. The use of synthetic data is due to a number of factors including:

1. The businesses contacted asked to keep the anonymity of their business data and transactions since the data used in this study would be open-source and uploaded to GitHub.
2. Since multiple businesses in the same sector were contacted to analyze data from, the business owners did not want their business data to leak to their competitors due to the open-source nature of this study.

2. METHODOLOGY

2.1 The Approach

This study follows a prototype-oriented design using simulated data and user interactions to model the functionality of an AI-driven business intelligence platform for small scale enterprises. The approach used was:

1. A dashboard was designed using ReactJS and Typescript via Vite to showcase business intelligence data (sales forecast, customer segmentation, and churn analysis).
2. Mock datasets were created to feed into machine learning models.
3. machine learning models were created using python libraries.
4. The models were deployed via a Fast API backend.

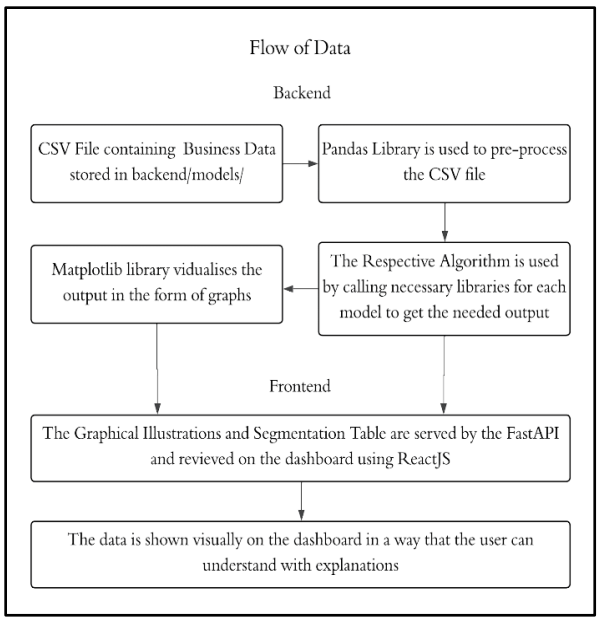
2.2 AI Model Development

Python Libraries Used: The scikit-learn library was utilized due to its extensive set of stable, thoroughly tested machine learning utilities, suited perfectly for quick prototyping of classification and clustering models during research. Sales forecasting utilized the XGBoost library due to its successful performance in processing tabular business data, resistance to missing information, as well as non-linear relationships better than standard regression models. The Pandas library was used for its efficiency in handling structured business datasets, enabling fast filtering, aggregation, and transformation of operational data into model-ready formats. NumPy library was utilized for its high-performance array manipulation and generation of random numbers, which were important for generating synthetic data with simulated noise for the purpose of simulation. The Matplotlib library was used due to its ability to create publication-quality, fully interactive plots ideal for academic papers.

Model 1: Sales Forecasting: A XGBoost Regressor Algorithm was used to forecast sales by taking date, product category, and historical sales volume as input. The output gives the predicted daily sales vs the actual sales data as a line graph.

Model 2: Customer Segmentation: A Kmeans Cluster Algorithm was used to segment customers by taking purchase frequency, average order value and recency as input. The output gives 4 scatter plots of Average Cart Value vs Total Money Spent highlighting customer behaviour across the various segments as and a table to understand the same.

Model 3: Churn Prediction: A DecisionTree Classifier Algorithm was used to predict the churn rate by taking last purchase, number of support tickets, and negative reviews as input. The output gives a line graph of Actual vs Predicted Churn according as a relation with customer satisfaction.



2.3 Ethical and Procedural Notes

All data used for the machine learning models were synthetically generated as discussed in the previous section. The results are framed accordingly, keeping in mind the synthetic nature of the output.

3. RESULTS AND DISCUSSIONS (PROTOTYPE SIMULATION)

3.1 Evaluation Metrics

The business intelligence data received from the machine learning models are evaluated on the following criterias:

1. The Mean Absolute Percentage Error (MAPE) [14] of the sales forecast to check the accuracy of the forecasting model in the simulated environment.
2. The Silhouette Score [15] of the customer segmentation model to ensure proper usability of the segmented information.
3. The Accuracy Score of the churn model to check the accuracy of the churn in the simulated environment.

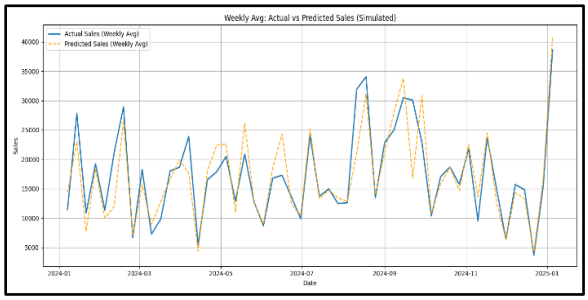
3.2 Model Performance Summary

Model	Algorithm Used	Simulated Metric	Outcome
Sales Forecasting	XGBoost Regressor	Mean Absolute % Error	18.7%

Customer Segmentation	KMeans (4 clusters)	Silhouette Score	0.52 (moderate separation)
Churn Prediction	DecisionTreeClassifier	Accuracy	84.3%

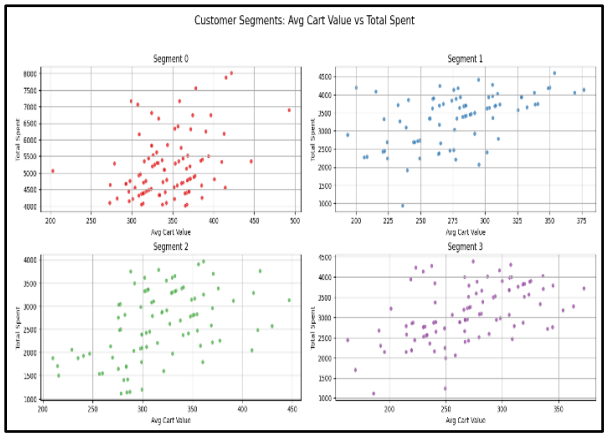
3.3 Visual Outputs

Sales Forecast Line Graph:



The visualization of sales predictions reveals how the model is capable of estimating underlying sales trends within weeks with relative steadiness. Despite existing slight deviations apparent from inherent noise, the forecasted curve does remarkably well echoing peaks and lows, showing the model's capability in estimating trends amidst inherent fluctuations in sales data.

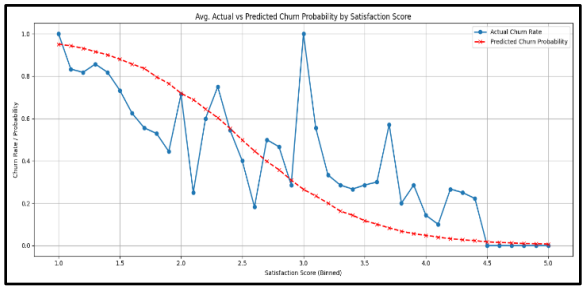
Customer Segmentation Scatter Plot:



Segment	Total Spent	Average Cart Value	Visits	Days Since Last Visit	Loyalty Score	Email Engagement %
0	5127.07	344.30	15	43	6.55	26.73
1	3354.39	277.99	12	61	8.02	42.75
2	2605.76	317.44	8	39	7.43	23.03
3	3092.54	268.67	12	30	4.93	28.28

The visualization of segmentation plots demonstrates behavioral clustering with KMeans (k=4). Segment 0 (top-left) identifies high spenders with high cart values as premium or bulk buyers. Segment 1 (top-right) indicates moderate spenders with equally distributed cart sizes. Segment 2 and Segment 3 (bottom) indicate low-spending customers with differently occurring visit frequencies. The distinction between plots confirm successful segmentation, offering targetable marketing strategies and tailored engagement according to spend behavior.

Churn Line Graph:



This visualization represents the interaction between customer satisfaction and churn probability. With increasing satisfaction levels, predicted churn (red line) goes down, simulating natural churn dynamics. Realized churn (blue line) varies sporadically, typical of inherent noise. Despite variables, the overall decreasing trend is as expected. This plot shows how predictive modeling can abstractly generalize patterns in churn even in simulated environments.

4. CONCLUSIONS

The research effectively built and tested a lightweight, module-based AI dashboard with synthetic business data, including forecasting, segmenting, and predicting churn modules that yielded usable outputs. The prototype illustrates low-cost AI applicability among small scale enterprises, even without large datasets, and acts as a model for wider implementation.

Success seen from the testing of these models shows the need for government- or incubator-sponsored deployment of similar structures in resource-limited areas.

Integration with accounting software like Tally or POS systems would add further versatility. From a research standpoint, future research must study AI preparedness matrices among small scale enterprises in Tier-II cities, monitor long-term behavioral changes

after adoption, and transfer the prototype into mobile-based dashboards for real-time information delivery.

ACKNOWLEDGMENT

I express special gratitude to Mrs. Soma Chakraborty, my Computer Science instructor, for her direction, encouragement, and guidance throughout this research. Her encouragement has gone a long way towards shaping both the technical and analytical aspect of this work. I am also grateful to the local entrepreneurs who offered informative insights and data patterns to encourage the simulations and practical implementation of this work. Their input served to ensure the work did not become isolated from the real-world challenges of their MSMEs. Finally, I appreciate the constant encouragement of peers and well-wishes who encouraged me to embark on this project with dedication and perseverance.

REFERENCE

[1] Press Information Bureau. (2018). "The Gross Value Added (GVA) by MSMEs in India's GDP was 29.7 % in 2017-18," https://www.pib.gov.in/PressReleaseDetailm.aspx?PRID=2087361

[2] My Republica. (2022). "Unlocking growth of small and medium-size enterprises," https://myrepublica.nagariknetwork.com/news/unlocking-growth-of-small-and-medium-size-enterprises/

[3] The Economic Times. (2025). "Digital Divide: MSMEs struggle to tap into AI revolution as high costs block access," https://economictimes.indiatimes.com/small-biz/sme-sector/digital-divide-msmes-struggle-to-tap-into-ai-revolution-as-high-costs-block-access/articleshow/121924294.cms

[4] Tornatzky, Louis G.; Fleischer, Mitchell (1990). The Processes of Technological Innovation. Issues in organization and management series. Lexington, Massachusetts: Lexington Books. ISBN 9780669203486. OCLC 20669819.

[5] Bagozzi, R. P., Davis, F. D., & Warshaw, P. R. (1992). Development and Test of a Theory of Technological Learning and Usage. Human Relations, 45(7), 659-686. https://doi.org/10.1177/001872679204500702

[6] J.D. Bryan, T. Zuva "A Review on TAM and TOE Framework Progression and How These Models Integrate", Advances in Science, Technology and Engineering Systems Journal, vol. 6, no. 3, pp. 137-145 (2021). https://www.astesj.com/v06/i03/p16/#1648228364276-af908526-b305

[7] Nurqamarani, A. S., Sogiarto, E., & Nurlaeli, N. (2021). Technology Adoption in Small-Medium Enterprises based on Technology Acceptance Model: A Critical Review. *Journal of Information Systems Engineering and Business Intelligence*, *7*(2), 162–172. https://doi.org/10.20473/jisebi.7.2.162-172

[8] Sadiq, Raghad & Safie, Zam & Abd Rahman, Abdul Hadi & Goudarzi, Shidrokh. (2021). Artificial intelligence maturity model: A systematic literature review. PeerJ Computer Science. 7. 10.7717/peerj-cs.661. https://www.researchgate.net/publication/354126381_Artificial_intelligence_maturity_model_a_systematic_literature_review

[9] Lichtenthaler, Ulrich. "Five Maturity Levels of Managing AI: From Isolated Ignorance to Integrated Intelligence." Journal of Innovation Management, 2020. https://journalsojs3.fe.up.pt/index.php/jim/article/view/2183-0606_008.001_0005

[10] Proietti, S., & Magnani, R. (2025). Assessing AI Adoption and Digitalization in SMEs: A Framework for Implementation. *ArXiv, abs/2501.08184*. https://arxiv.org/pdf/2501.08184

[11] Yusuf, Samuel & Durodola, Remilekun & Ocran, Godbless & Abubakar, Justina & Amarachi, Zita & Echere, Amarachi & Paul-Adeleye, Adedamola. (2024). Challenges and opportunities in AI and digital transformation for SMEs: A cross- continental perspective. World Journal of Advanced Research and Reviews. Volume 23. 668-678. 10.30574/wjarr.2024.23.3.2511.https://wjarr.com/content/challenges-and-opportunities-ai-and-digital-transformation-smes-cross-continental

[12] Agostino, Deborah & Arena, Marika & Azzone, Giovanni & Dal Molin, Martina & Masella, Cristina. (2012). Developing a performance measurement system for public research centres. International Journal of Business Science and

Applied Management. 7. 40-60. 10.69864/ijbsam.7-1.81. https://www.business-and-management.org/library/2012/7_1--43-60-Agostino,Arena,Azzone,Molin,Masella.pdf

[13] Myttenaere, A.D., Golden, B., Grand, B.L., & Rossi, F. (2016). Mean Absolute Percentage Error for regression models. Neurocomputing, 192, 38-48. https://arxiv.org/pdf/1605.02541

[14] Peter J. Rousseeuw (1987). "Silhouettes: a Graphical Aid to the Interpretation and Validation of Cluster Analysis". Computational and Applied Mathematics. 20: 53–65. doi:10.1016/0377-0427(87)90125-7.

Additional Resources Used

Books:

[1] Rick Sherman. 2014. Business Intelligence Guidebook: From Data Integration to Analytics (1st. ed.). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA. Retrieved from: https://www.educate.elsevier.com/book/details/9780124114616

[2] Ian Goodfellow, Yoshua Bengio, and Aaron Courville 2016. *Deep Learning*. MIT Press. Retrieved from: https://books.google.co.in/books?id=Np9SDQAAQBAJ&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false

[3] Geron, A. 2019. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media, Inc. Retrieved from: https://www.amazon.com/Hands-Machine-Learning-Scikit-Learn-TensorFlow/dp/1492032646

[4] Krug, S. 2005. *Don't Make Me Think: A Common Sense Approach to the Web (2nd Edition)*. New Riders Publishing. Retrieved from: https://dn790002.ca.archive.org/0/items/SteveKrugDontMakeMeThink/Steve_Krug_Don%E2%80%99t_Make_Me_Think%2C.pdf

Dictionaries:

[1] Cambridge Dictionary. Retrieved from: https://dictionary.cambridge.org/us/dictionary/english/

GitHub Repository of the Prototype:

[1] https://github.com/SusmitAcharya/nivritti_ai