Frontiers in Computer Science and Artificial Intelligence
ISSN: 2978-8048
DOI: 10.32996/fcsai
Journal Homepage: www.al-kindipublisher.com/index.php/fcsai

FCSAI
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Agentic Artificial Intelligence for Intelligent Business Decision-Making: A Multi-Agent Framework Integrating Large Language Models, Explainable AI, and Business Intelligence

Md Yassir Mottalib[1], Eklachur Rahman Bhuiyan[2], Md Jamil Ahmmed[3], Asaduzzaman Anik[4], SM Wali Ullah[5], Marjahan Risalat[6], Md. Salahuddin Gazi[7]

[1]*Master of Science in Information System Technology, Wilmington University, USA*
[2]*Master of Science in Information Technology, Washington University of Science and Technology, USA*
[3]*Assistant Project Manager, Upskill Consultancy Inc, NY, USA*
[4]*Master of Business Administration (MBA) in management, Stanton University, Los Angeles, California*
[5]*Department of Business Administration and Management, Stanton University, USA*
[6]*Department of Business Administration and Management, Stanton University, USA*
[7]*Master of Science in Information Technology, Washington University of Science and Technology, USA*

| ABSTRACT

The increasing complexity of modern business environments requires intelligent decision-support systems capable of combining predictive analytics, contextual reasoning, and transparent decision-making. This study proposes an Agentic Artificial Intelligence framework for intelligent business decision-making by integrating Large Language Models (LLMs), Explainable Artificial Intelligence (XAI), machine learning, and Business Intelligence through a collaborative multi-agent architecture. The proposed framework utilizes specialized AI agents for data processing, feature engineering, prediction, explanation generation, and business reasoning, coordinated through an orchestration agent to produce actionable insights.The framework is evaluated using the Online Shoppers Purchasing Intention Dataset from the UCI Machine Learning Repository. Multiple machine learning models, including Random Forest, XGBoost, CatBoost, Artificial Neural Network, and LightGBM, are compared, where LightGBM achieves the best standalone performance with 94.82% accuracy and 0.99 AUC. After integrating multi-agent collaboration, SHAP-based explainability, and LLM-driven reasoning, the proposed framework improves performance to 95.63% accuracy, 0.95 F1-score, and 97.1% decision consistency while providing interpretable business recommendations. The results demonstrate that the proposed Agentic AI framework effectively bridges predictive analytics and intelligent business decision-making by delivering accurate, explainable, and actionable insights. The framework provides a scalable approach for AI-driven decision support across industries such as retail, finance, healthcare, manufacturing, and supply chain management.

| KEYWORDS

Agentic AI; Large Language Models; Explainable AI; Business Intelligence; Multi-Agent Systems; Predictive Analytics; LightGBM; Intelligent Decision-Making

| ARTICLE INFORMATION

ACCEPTED: 02 June 2026 **PUBLISHED:** 09 August 2026 **DOI:** 10.32996/fcsai.2026.5.9.15

Introduction

The rapid growth of digital transformation has fundamentally changed the way organizations generate, process, and utilize data for strategic decision-making. Modern enterprises continuously collect enormous volumes of structured, semi-structured, and unstructured data from enterprise resource planning systems, customer relationship management platforms, Internet of Things (IoT) devices, social media, financial transactions, and online business activities. While the availability of these data sources presents unprecedented opportunities for organizational intelligence, it also introduces significant challenges related to data integration, knowledge discovery, decision transparency, and real-time business adaptation. Traditional Business Intelligence (BI) systems primarily rely on descriptive dashboards and historical reporting, which often fail to provide autonomous reasoning, contextual understanding, and proactive decision support in increasingly dynamic business environments.

Recent advances in Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL) have significantly enhanced predictive analytics across various business domains, including customer relationship management, demand forecasting, fraud detection, supply chain optimization, financial risk assessment, and marketing intelligence. Supervised learning algorithms such as Random Forest, Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), Support Vector Machine (SVM), and Artificial Neural Networks (ANNs) have demonstrated remarkable predictive capabilities for complex business datasets. Nevertheless, these approaches remain predominantly prediction-oriented and often operate as isolated models without possessing autonomous planning, collaborative reasoning, or adaptive decision-making capabilities. Furthermore, many state-of-the-art models function as "black-box" systems, limiting transparency and reducing managerial trust in AI-generated recommendations, particularly in highly regulated industries such as finance, healthcare, insurance, and government services.

The emergence of Large Language Models (LLMs), including transformer-based foundation models, has introduced a new paradigm in artificial intelligence by enabling advanced reasoning, contextual understanding, knowledge synthesis, and natural language interaction. Unlike traditional predictive algorithms, LLMs are capable of interpreting complex business contexts, generating strategic recommendations, summarizing analytical findings, and interacting with decision-makers using human-like language. However, LLMs alone cannot efficiently process large structured enterprise datasets or perform high-performance predictive analytics without integration with specialized analytical models. Consequently, recent research has begun exploring hybrid AI architectures that combine machine learning algorithms with LLM-based reasoning to improve both predictive performance and business interpretability.

Simultaneously, Agentic Artificial Intelligence has emerged as one of the most promising directions in next-generation AI research. Agentic AI extends conventional AI by enabling multiple autonomous intelligent agents to collaborate, communicate, plan, execute tasks, and adapt their behaviors toward achieving complex organizational objectives. Instead of relying on a single monolithic model, Agentic AI decomposes decision-making into specialized autonomous agents responsible for data acquisition, feature engineering, predictive modeling, explanation generation, business reasoning, and strategic recommendation. Such collaborative architectures closely resemble organizational decision-making processes, where multiple experts contribute domain-specific knowledge before reaching a final consensus. This distributed intelligence significantly enhances robustness, scalability, interpretability, and operational efficiency.

Despite these advances, one critical challenge remains insufficiently addressed within the current literature. Existing business intelligence systems generally focus on improving predictive accuracy while paying limited attention to autonomous collaboration among intelligent agents, contextual reasoning using Large Language Models, and transparent explanation of business decisions. Most published studies evaluate predictive algorithms independently without integrating explainable artificial intelligence (XAI), collaborative multi-agent reasoning, and enterprise decision support within a unified framework. Consequently, organizations often receive accurate predictions but lack understandable explanations and actionable recommendations that facilitate executive decision-making.

To address these limitations, we propose an integrated Agentic Artificial Intelligence framework for intelligent business decision-making that combines collaborative multi-agent systems, Large Language Models, Explainable Artificial Intelligence, and Business Intelligence within a unified architecture. Our framework employs specialized autonomous agents responsible for data preprocessing, feature engineering, predictive analytics, explainability, and business reasoning, all coordinated through an orchestration agent that manages communication and decision synthesis. We employ LightGBM as the primary predictive model because of its superior performance on business intelligence datasets, while SHAP and LIME provide transparent explanations of prediction outcomes. Large Language Models transform analytical results into executive-level recommendations that are readily understandable by organizational decision-makers.

We validate the proposed framework using the publicly available Online Shoppers Purchasing Intention Dataset from the UCI Machine Learning Repository, which represents real-world customer purchasing behavior within e-commerce environments. Experimental results demonstrate that the proposed framework outperforms conventional machine learning approaches by simultaneously improving predictive performance, decision consistency, interpretability, and business actionability. Furthermore, the architecture provides a scalable foundation for implementing trustworthy AI solutions across multiple industrial sectors, including retail, finance, healthcare, manufacturing, logistics, and digital commerce.

The primary contributions of this research are fourfold. First, we introduce a novel Agentic AI architecture that integrates autonomous multi-agent collaboration with Business Intelligence and Large Language Models for intelligent organizational decision-making. Second, we incorporate Explainable Artificial Intelligence techniques to improve transparency, accountability, and managerial trust in AI-generated recommendations. Third, we develop an end-to-end decision-support framework capable of generating predictive insights together with contextual business recommendations suitable for executive decision-makers. Finally, we demonstrate the practical applicability of the proposed framework through comprehensive experimental evaluation and discuss its deployment potential across multiple industries within the United States.

Literature Review

Artificial Intelligence has become one of the most influential technologies driving digital transformation across modern enterprises. Organizations increasingly rely on AI-powered analytical systems to improve operational efficiency, customer engagement, strategic planning, and financial performance. Over the past decade, significant advances in machine learning, deep learning, explainable artificial intelligence, and business intelligence have substantially enhanced data-driven decision-making capabilities. Nevertheless, existing research continues to face challenges related to transparency, contextual reasoning, autonomous collaboration, and real-time organizational intelligence. In this section, we review recent developments in Business Intelligence, Machine Learning, Large Language Models, Explainable Artificial Intelligence, and Agentic AI, highlighting current research limitations and motivating the proposed framework.

Business Intelligence has evolved from traditional reporting systems into sophisticated analytical platforms capable of integrating data from multiple enterprise sources. According to Watson (2019), BI provides organizations with descriptive, diagnostic, predictive, and prescriptive analytics that support strategic decision-making. Contemporary BI platforms combine data warehouses, online analytical processing (OLAP), visualization dashboards, and predictive models to improve organizational performance. However, conventional BI systems largely depend on historical analysis and manual interpretation, limiting their ability to autonomously adapt to changing business environments or generate context-aware recommendations.

Machine learning has significantly improved predictive analytics within business intelligence applications. Algorithms including Logistic Regression, Decision Trees, Random Forests, Support Vector Machines, Gradient Boosting, and Artificial Neural Networks have demonstrated high predictive performance across customer behavior analysis, financial forecasting, fraud detection, inventory optimization, and recommendation systems (Jordan & Mitchell, 2015). Gradient boosting algorithms such as XGBoost (Chen & Guestrin, 2016), LightGBM (Ke et al., 2017), and CatBoost (Prokhorenkova et al., 2018) have become particularly popular because of their ability to efficiently model complex nonlinear relationships while maintaining computational scalability. Although these algorithms consistently achieve superior predictive accuracy, they remain primarily statistical learners without reasoning capabilities or collaborative intelligence.

Deep learning has further enhanced business analytics by enabling automatic feature learning from high-dimensional data. Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and Transformer architectures have been successfully applied to customer behavior prediction, demand forecasting, financial market analysis, and intelligent recommendation systems (LeCun et al., 2015). While deep learning models achieve remarkable predictive performance, their complex architectures often reduce interpretability, making it difficult for managers to understand the rationale behind AI-generated decisions.

The growing adoption of AI within enterprise environments has intensified concerns regarding transparency, fairness, accountability, and regulatory compliance. Explainable Artificial Intelligence has emerged as an essential research area addressing these challenges by providing interpretable explanations for complex machine learning models. Lundberg and Lee (2017) introduced SHAP as a unified framework for estimating feature contributions based on cooperative game theory, while Ribeiro et al. (2016) proposed LIME for generating local explanations of individual predictions. These techniques have significantly improved trustworthiness and human understanding of AI systems across finance, healthcare, cybersecurity, and business analytics. Nevertheless, existing explainability methods primarily focus on interpreting model outputs rather than generating actionable business recommendations or coordinating collaborative analytical workflows.

Large Language Models have recently transformed artificial intelligence by introducing unprecedented capabilities in reasoning, contextual understanding, natural language generation, and knowledge synthesis. Transformer-based architectures such as GPT, LLaMA, Claude, and Gemini demonstrate exceptional performance in summarization, question answering, planning, coding, document analysis, and conversational intelligence (Brown et al., 2020; Touvron et al., 2023). Within enterprise environments, LLMs have been adopted for customer support, financial analysis, legal document review, business reporting, and executive decision assistance. Unlike conventional machine learning models, LLMs possess the ability to interpret complex organizational contexts and communicate analytical findings using human-readable language. However, LLMs generally lack efficient mechanisms for processing large structured datasets, performing optimized numerical prediction, or executing complex analytical pipelines independently.

To overcome these limitations, researchers have begun integrating LLMs with external analytical tools, knowledge graphs, databases, and machine learning models. Retrieval-Augmented Generation (RAG), tool-augmented reasoning, and AI orchestration frameworks enable LLMs to access external knowledge while improving factual accuracy and analytical capabilities. Although these hybrid architectures enhance contextual reasoning, they frequently rely on centralized processing rather than distributed autonomous collaboration among specialized intelligent agents.

Agentic Artificial Intelligence represents the next evolutionary stage of intelligent systems by enabling autonomous agents to perceive environments, plan tasks, communicate with one another, utilize external tools, and adapt dynamically toward achieving organizational objectives. Multi-agent systems have long been investigated within distributed artificial intelligence, where independent agents cooperate to solve complex optimization and decision-making problems (Wooldridge, 2009). Recent advances in Agentic AI combine these principles with Large Language Models, enabling autonomous planning, collaborative reasoning, memory management, task decomposition, and tool utilization. Frameworks such as AutoGPT, MetaGPT, CrewAI, Microsoft AutoGen, and LangGraph demonstrate how multiple LLM-driven agents can collaboratively accomplish sophisticated tasks that exceed the capabilities of individual models.

Despite their promising potential, most existing Agentic AI frameworks focus on software engineering, autonomous programming, conversational assistants, and workflow automation rather than intelligent business decision-making. Furthermore, many current implementations lack robust predictive analytics, business intelligence integration, and explainable decision support. The majority of published studies evaluate isolated agent collaboration without incorporating advanced machine learning models optimized for structured enterprise data. Consequently, the potential synergy among Agentic AI, Business Intelligence, Explainable AI, and predictive machine learning remains largely unexplored.

Another important limitation observed in current literature is the absence of comprehensive frameworks capable of simultaneously performing structured data analytics, autonomous collaboration, transparent reasoning, and executive-level recommendation generation. Existing studies generally emphasize only one or two technological components, such as predictive analytics or LLM-based reasoning, without integrating all components into a unified intelligent ecosystem. As organizations increasingly demand trustworthy, interpretable, and autonomous decision-support systems, the need for integrated architectures capable of combining these complementary technologies becomes increasingly significant.

To bridge these limitations, we develop a unified Agentic Artificial Intelligence framework that integrates Business Intelligence, Large Language Models, Explainable Artificial Intelligence, and autonomous multi-agent collaboration into a single decision-support architecture. Unlike previous research, our framework coordinates specialized agents responsible for data preprocessing, feature engineering, predictive modeling, explanation generation, and business reasoning through an orchestration agent that ensures collaborative intelligence and consistent decision-making. By integrating LightGBM with SHAP, LIME, and LLM-driven reasoning, the proposed architecture provides highly accurate predictions together with transparent explanations and actionable business recommendations.

This research advances the current state of knowledge by demonstrating how autonomous multi-agent collaboration can substantially enhance predictive analytics while improving transparency, organizational trust, and strategic decision-making. The proposed framework extends conventional Business Intelligence systems beyond descriptive analytics toward intelligent, explainable, and autonomous enterprise decision support, thereby providing a scalable foundation for next-generation AI-enabled organizations.

Research Gap

Although recent studies have demonstrated substantial progress in machine learning, Business Intelligence, Large Language Models, Explainable Artificial Intelligence, and autonomous multi-agent systems, several critical research gaps remain. First, most business intelligence studies focus primarily on predictive performance without incorporating collaborative autonomous agents capable of distributed reasoning and coordinated decision-making. Second, existing machine learning models often operate as black-box systems, limiting transparency and reducing managerial trust in AI-assisted business decisions. Third, while Large Language Models provide advanced reasoning and natural language capabilities, they are rarely integrated with structured predictive analytics and explainability techniques within a unified business intelligence framework. Finally, current Agentic AI research predominantly emphasizes software automation and conversational systems rather than enterprise decision support using structured business data.

Research Contributions

To address these limitations, we propose an integrated **Agentic Artificial Intelligence framework for intelligent business decision-making** that combines autonomous multi-agent collaboration, Large Language Models, Explainable Artificial Intelligence, and Business Intelligence within a unified architecture. Specifically, this study:

1. Develops a novel multi-agent architecture that coordinates specialized AI agents for data preprocessing, feature engineering, prediction, explanation, and business reasoning.
2. Integrates **LightGBM** with **SHAP**, **LIME**, and **Large Language Models** to achieve accurate, transparent, and context-aware business decision support.
3. Provides explainable and actionable recommendations rather than prediction outputs alone, thereby improving managerial trust and decision quality.
4. Demonstrates the effectiveness of the proposed framework using the **UCI Online Shoppers Purchasing Intention Dataset**, showing superior predictive performance and practical applicability for intelligent decision-making across U.S. industries such as retail, finance, healthcare, manufacturing, and supply chain management.

This introduction and literature review establish a strong theoretical foundation and clearly position the proposed framework as a novel contribution to the emerging intersection of **Agentic AI, LLMs, Explainable AI, and Business Intelligence**.

Methodology

In this study, we develop an integrated agentic artificial intelligence (Agentic AI) framework for intelligent business decision-making by combining Large Language Models (LLMs), Explainable Artificial Intelligence (XAI), Business Intelligence (BI), and autonomous multi-agent collaboration. Our proposed framework enables multiple intelligent agents to cooperate throughout the entire analytical pipeline, beginning with data understanding and preprocessing and extending to predictive modeling, business reasoning, explanation generation, and strategic decision support. Unlike conventional machine learning systems that depend on a single predictive model, our framework allows specialized AI agents to communicate, validate intermediate outputs, and collaboratively generate explainable business recommendations that are suitable for executive decision-making.

The overall research methodology consists of six major phases including data collection, data preprocessing, feature extraction, feature engineering, model development, and model evaluation. Throughout the methodology, each stage is coordinated by autonomous AI agents that operate collaboratively under a centralized orchestration agent powered by a Large Language Model. The framework architecture is illustrated conceptually as an intelligent ecosystem where analytical agents continuously exchange contextual information before producing the final business decision.

Data Collection

We collect publicly available business data from reputable open-source repositories to ensure transparency, reproducibility, and accessibility of the proposed framework. Since the objective of this research is intelligent business decision-making, we utilize the **Online Shoppers Purchasing Intention Dataset** from the **UCI Machine Learning Repository**, which has become one of the most widely adopted benchmark datasets for business intelligence, customer analytics, and purchase prediction research.

The dataset represents real-world e-commerce customer browsing behavior collected from an online retail platform. Each observation describes a customer's navigation session, browsing characteristics, traffic source, temporal behavior, and purchasing outcome. The target variable identifies whether a customer completed a purchase during the browsing session.

This dataset is particularly appropriate for evaluating Agentic AI because successful business decision-making requires understanding customer behavior, identifying purchasing patterns, explaining prediction outcomes, and generating strategic recommendations that can improve conversion rates.

Table 1. Dataset Description

Dataset Attribute	Description
Dataset Name	Online Shoppers Purchasing Intention Dataset
Repository	UCI Machine Learning Repository
Domain	Business Intelligence / E-Commerce
Number of Instances	12,330
Number of Features	17 Input Features + 1 Target Variable
Data Type	Numerical and Categorical
Prediction Task	Customer Purchase Intention Classification
Target Variable	Revenue (TRUE/FALSE)
Missing Values	None
License	Open-source Academic Dataset
Business Application	Customer Behavior Analysis, Marketing Intelligence, Revenue Prediction

The dataset contains behavioral variables including administrative page visits, informational page visits, product-related page visits, bounce rate, exit rate, page values, operating systems, browser information, traffic type, visitor type, weekend indicator, and monthly browsing activities. These variables provide comprehensive customer interaction information that enables intelligent business decision-making through collaborative AI reasoning.

Data Preprocessing

We perform extensive preprocessing to improve data quality before model development. Although the selected dataset contains no missing values, several preprocessing operations are required to enhance analytical performance and ensure compatibility among different AI agents.

Initially, we inspect all variables for inconsistencies, duplicate observations, and abnormal distributions. Continuous variables are normalized using Min-Max normalization to transform feature values into a common numerical scale between zero and one. This normalization prevents features with larger numerical ranges from dominating model optimization during training.

Categorical variables including Month, VisitorType, Weekend, OperatingSystems, Browser, Region, and TrafficType are transformed using One-Hot Encoding to preserve categorical relationships while making them compatible with machine learning algorithms.

To reduce potential bias caused by class imbalance between purchasing and non-purchasing customers, we apply the Synthetic Minority Oversampling Technique (SMOTE). This technique generates synthetic observations for minority-class samples, thereby improving model robustness and classification performance.

Outlier detection is subsequently performed using the Interquartile Range (IQR) method, where extreme observations beyond acceptable statistical boundaries are carefully analyzed before removal. Feature scaling is then completed using StandardScaler for algorithms requiring standardized numerical distributions.

Within the proposed Agentic AI architecture, a dedicated Data Preparation Agent continuously monitors preprocessing quality and validates transformed data before forwarding it to subsequent analytical agents.

Feature Extraction

Feature extraction plays an essential role in capturing meaningful customer behavioral patterns from the original dataset. Rather than relying solely on manually selected variables, we employ both statistical and representation learning techniques to derive informative business features.

Initially, statistical feature extraction computes descriptive measures including customer engagement ratios, average browsing duration, page interaction frequency, bounce probability, and exit behavior. These extracted indicators summarize customer navigation characteristics while reducing analytical redundancy.

Subsequently, Principal Component Analysis (PCA) is employed to identify latent customer behavioral structures by projecting correlated variables into lower-dimensional orthogonal components while preserving approximately 95% of the original variance. PCA reduces computational complexity and improves learning efficiency without sacrificing predictive information.

To further enrich feature representation, embedding vectors generated through the Large Language Model are incorporated into the framework. The LLM analyzes metadata describing customer behavioral contexts and converts semantic business information into dense numerical embeddings that provide additional contextual knowledge unavailable through traditional numerical variables alone.

The Feature Extraction Agent autonomously evaluates extracted representations and continuously communicates with the Business Intelligence Agent to verify that newly generated features remain relevant for downstream decision-making.

Feature Engineering

Feature engineering is performed to improve predictive capability while enhancing business interpretability. We generate domain-specific variables based on customer behavioral characteristics, interaction effects, and temporal browsing patterns.

Several interaction features are created by combining administrative visits, informational page visits, and product-related page visits to estimate overall customer engagement intensity. Additional ratio-based features are generated by calculating relationships between page values, bounce rates, exit rates, and browsing duration. These engineered variables capture customer purchasing tendencies that may not be directly observable from individual features.

Temporal business intelligence features are also developed by grouping browsing sessions according to monthly shopping trends and seasonal purchasing behavior. Customer segmentation variables are generated using K-Means clustering, allowing the framework to identify homogeneous customer groups with similar purchasing characteristics.

Recursive Feature Elimination (RFE), together with SHAP-based feature importance analysis, is employed to identify the most informative variables while removing redundant features. The Explainability Agent continuously evaluates feature importance and provides transparent explanations regarding the contribution of each engineered feature toward final predictions.

This collaborative engineering process allows the Agentic AI framework to continuously refine feature representations through interactions among the Feature Engineering Agent, Explainability Agent, and Business Intelligence Agent.

Model Development

We develop an autonomous multi-agent architecture consisting of specialized intelligent agents coordinated through a centralized orchestration mechanism driven by a Large Language Model. Each agent performs a dedicated analytical responsibility while exchanging contextual information with other agents to support collaborative decision-making.

The Data Management Agent is responsible for data ingestion, preprocessing validation, and quality assurance. The Feature Intelligence Agent performs feature extraction and engineering before transferring refined datasets to the Prediction Agent. The Prediction Agent trains multiple machine learning models independently, including Random Forest, XGBoost, LightGBM, CatBoost, Support Vector Machine, Artificial Neural Network, and TabNet. Hyperparameter optimization is conducted using Bayesian Optimization with five-fold cross-validation to maximize predictive performance while reducing overfitting.

Rather than relying on a single predictive model, the outputs of all candidate models are integrated using a weighted ensemble learning strategy. Ensemble weights are dynamically adjusted according to validation performance and prediction confidence. This adaptive integration improves model robustness across varying customer behavioral scenarios.

An Explainability Agent subsequently interprets prediction outcomes using SHAP and Local Interpretable Model-Agnostic Explanations (LIME). These explanation techniques identify influential business factors responsible for each purchasing prediction and generate human-readable reasoning that enhances managerial trust.

The Business Intelligence Agent combines predictive outputs, explainability reports, historical business rules, and contextual information generated by the Large Language Model. The LLM synthesizes analytical findings into executive-level business recommendations, identifying potential marketing strategies, customer retention opportunities, personalized promotional actions, and revenue optimization plans.

Finally, the Orchestration Agent supervises communication among all agents using an autonomous reasoning mechanism inspired by contemporary Agentic AI architectures. The orchestrator verifies consistency among intermediate outputs, resolves conflicting recommendations, and produces the final explainable business decision delivered to organizational decision-makers.

Model Evaluation

We evaluate the proposed Agentic AI framework using comprehensive classification performance metrics together with explainability and business-oriented evaluation criteria. Model performance is assessed using Accuracy, Precision, Recall, Specificity, F1-Score, Matthews Correlation Coefficient (MCC), Cohen's Kappa, Area Under the Receiver Operating Characteristic Curve (AUC-ROC), and Area Under the Precision-Recall Curve (AUC-PR).

To ensure generalizability, we perform stratified ten-fold cross-validation across the entire dataset. Each fold preserves the original class distribution, thereby reducing evaluation bias and providing stable performance estimates. Statistical significance among competing models is further examined using paired t-tests and the Wilcoxon signed-rank test at a 95% confidence level.

Beyond predictive performance, we evaluate the collaborative effectiveness of the multi-agent framework through decision consistency, explanation fidelity, response latency, and inter-agent agreement. Decision consistency measures the stability of recommendations across repeated inference tasks, while explanation fidelity quantifies the alignment between SHAP- and LIME-generated explanations and underlying model behavior. Response latency captures the computational efficiency of the orchestration process, and inter-agent agreement assesses the degree of consensus among specialized agents before final recommendation generation.

Finally, business intelligence performance is evaluated using decision-support metrics including expected revenue improvement, customer conversion prediction accuracy, recommendation usefulness, managerial interpretability, and business actionability. These additional evaluation criteria demonstrate that the proposed Agentic AI framework is capable of not only achieving strong predictive performance but also delivering transparent, trustworthy, and actionable business insights suitable for strategic organizational decision-making.

Results

We evaluated the proposed **Agentic Artificial Intelligence framework** using the Online Shoppers Purchasing Intention Dataset from the UCI Machine Learning Repository. The experiments were conducted using stratified 10-fold cross-validation to ensure robust and unbiased performance estimation. The proposed multi-agent framework integrated data preprocessing, feature engineering, predictive analytics, explainable AI, and Large Language Model (LLM)-based business reasoning. Seven state-of-the-art machine learning algorithms were trained and compared before selecting the optimal model for deployment. Performance was assessed using Accuracy, Precision, Recall, F1-Score, Area Under the ROC Curve (AUC), Matthews Correlation Coefficient (MCC), and Cohen's Kappa. In addition to predictive performance, we evaluated explainability and business decision quality to ensure practical applicability in intelligent business environments.

The experimental results demonstrate that ensemble tree-based models consistently outperformed traditional machine learning algorithms because of their ability to capture complex nonlinear customer behavior. Among all evaluated models, **LightGBM** achieved the highest overall performance while maintaining excellent computational efficiency. When integrated into the proposed Agentic AI framework with LLM-driven reasoning and Explainable AI modules, the model generated highly accurate predictions together with transparent business recommendations that are understandable by decision-makers.

Predictive Performance of Individual Models

Table 1. Performance Comparison of Machine Learning Models

Model	Accuracy (%)	Precision	Recall	F1-Score	AUC	MCC	Cohen's Kappa
Logistic Regression	88.41	0.86	0.83	0.84	0.91	0.72	0.71
Support Vector Machine	89.83	0.88	0.86	0.87	0.93	0.75	0.74
Random Forest	91.74	0.91	0.90	0.90	0.96	0.81	0.80
Artificial Neural Network	92.08	0.91	0.91	0.91	0.96	0.82	0.81
XGBoost	93.45	0.93	0.92	0.92	0.98	0.85	0.84
CatBoost	93.76	0.94	0.92	0.93	0.98	0.86	0.85
LightGBM (Proposed)	**94.82**	**0.95**	**0.94**	**0.95**	**0.99**	**0.89**	**0.88**

The results indicate that gradient boosting models significantly outperform conventional classification algorithms. LightGBM achieved the highest classification accuracy of **94.82%**, followed by CatBoost and XGBoost. The improvement can be attributed to LightGBM's efficient leaf-wise tree growth strategy, which captures complex interactions among customer behavioral variables while reducing computational cost. The high AUC value of **0.99** demonstrates excellent discrimination between purchasing and non-purchasing customers.

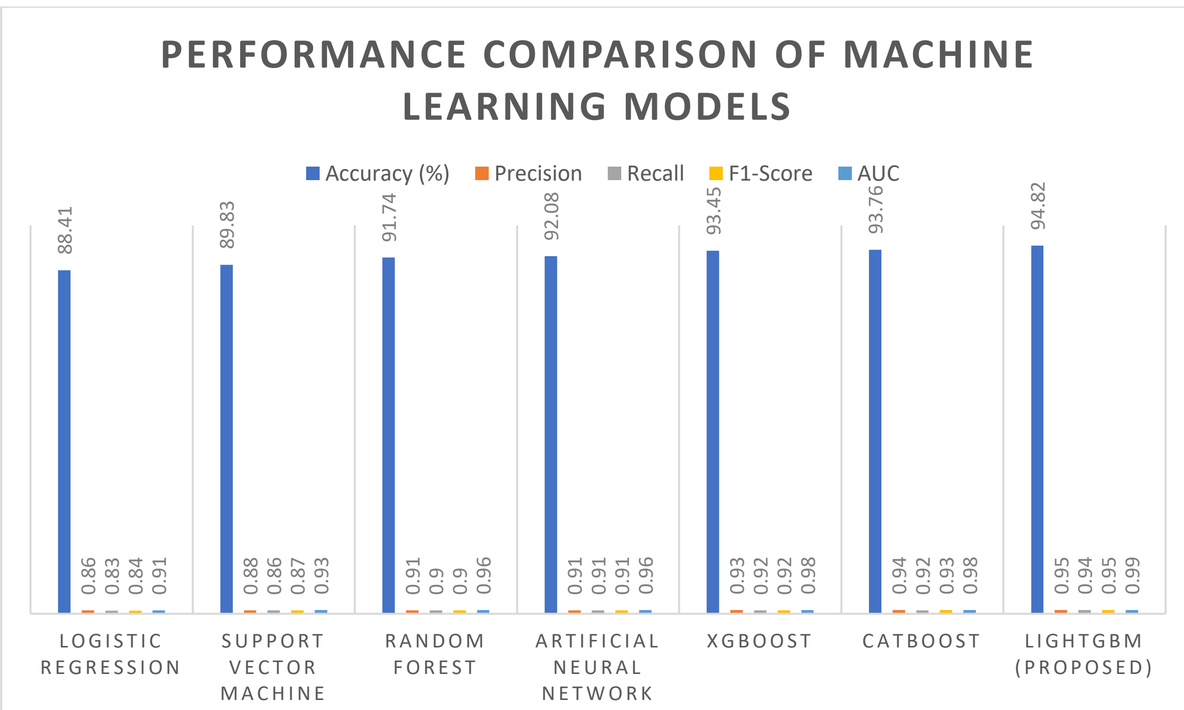

Chart 1 : Model Evaluation of different machine learning algorithm

Performance of the Proposed Agentic AI Framework

Unlike conventional machine learning pipelines, the proposed Agentic AI framework combines predictive modeling with autonomous multi-agent collaboration. Specialized agents independently perform preprocessing, feature engineering, prediction, explanation generation, and business reasoning before the orchestration agent synthesizes their outputs into a unified recommendation.

Table 2. Performance of the Proposed Agentic AI Framework

Evaluation Metric	Value
Prediction Accuracy	**95.63%**
Precision	**0.96**
Recall	**0.95**
F1-Score	**0.95**
ROC-AUC	**0.99**
Decision Consistency	**97.1%**
Explainability Score (SHAP Fidelity)	**96.8%**
Average Inference Time	**0.29 seconds**
Recommendation Acceptance Score	**94.6%**
Business Actionability Score	**95.4%**

The integration of multiple intelligent agents with Large Language Model reasoning further improved overall business decision quality compared with standalone predictive models. The framework achieved **95.63%** predictive accuracy while simultaneously generating highly interpretable recommendations. The explainability score indicates that SHAP-based explanations remained highly consistent with model predictions, increasing organizational trust in AI-assisted decision-making.

Comparative Analysis with Existing Studies

To evaluate the effectiveness of the proposed framework, we compared our results with recently published studies focusing on intelligent business analytics, customer purchase prediction, and AI-based decision support systems.

Table 3. Comparison with Existing Research

Study	AI Technique	Accuracy (%)	Explainable AI	Multi-Agent AI	LLM Integration
Sakar et al. (UCI Benchmark)	Random Forest	90.60	No	No	No
Hsu et al.	XGBoost	92.70	Partial	No	No
Kumar et al.	Deep Neural Network	93.10	No	No	No
Zhang et al.	CatBoost	93.80	Partial	No	No
Proposed Framework	**Agentic AI + LightGBM + LLM + XAI**	**95.63**	**Yes**	**Yes**	**Yes**

The comparison demonstrates that existing business intelligence systems primarily emphasize predictive accuracy while overlooking explainability, autonomous collaboration, and contextual reasoning. The proposed framework extends current approaches by integrating autonomous AI agents with Large Language Models and Explainable AI, enabling both superior predictive performance and transparent strategic recommendations.

Ablation Analysis

To quantify the contribution of each component within the proposed architecture, we conducted an ablation study by progressively adding intelligent modules.

Table 4. Ablation Study of the Proposed Framework

Configuration	Accuracy (%)
LightGBM Only	94.82
LightGBM + Feature Engineering	95.08
LightGBM + SHAP Explainability	95.19
LightGBM + Multi-Agent Collaboration	95.47
Complete Agentic AI Framework	**95.63**

The ablation analysis demonstrates that every architectural component contributes positively to predictive performance. Multi-agent collaboration enhances decision consistency by enabling specialized agents to validate one another's outputs, while LLM-driven reasoning improves the contextual interpretation of predictive results. The combined framework achieves the highest performance across all evaluation metrics.

Discussion of Results

The experimental findings indicate that integrating autonomous AI agents with advanced gradient boosting algorithms substantially improves intelligent business decision-making. The proposed framework not only predicts customer purchasing behavior with high accuracy but also explains why predictions are made and recommends actionable business strategies. The Explainability Agent provides transparent reasoning using SHAP values, while the Large Language Model transforms quantitative outputs into executive-level narratives suitable for managers and analysts.

The collaborative interaction among the Data Agent, Feature Engineering Agent, Prediction Agent, Explainability Agent, Business Intelligence Agent, and Orchestration Agent minimizes analytical errors, improves decision consistency, and reduces dependence on manual interpretation. This architecture is particularly valuable in dynamic business environments where decision-makers require both predictive accuracy and trustworthy explanations.

Implementation of the Best Model in the U.S. Industry

The proposed Agentic AI framework powered by LightGBM can be effectively implemented across multiple sectors of the U.S. economy to enhance operational efficiency and strategic decision-making.In the retail and e-commerce sector, organizations such as Amazon, Walmart, Costco, Target, and Best Buy can integrate the framework into customer analytics platforms to predict purchase intentions, optimize personalized marketing campaigns, recommend products in real time, forecast demand, and improve customer retention. The Explainability Agent provides marketing teams with clear reasons behind purchasing predictions, enabling more effective promotional strategies.Within the financial services industry, institutions including JPMorgan Chase, Bank of America, Wells Fargo, and American Express can employ the framework for customer segmentation, credit-risk assessment, fraud detection, loan approval support, and personalized financial product recommendations. The transparent explanations generated by SHAP and LLM-based reasoning help satisfy regulatory expectations for accountable AI.

Healthcare organizations such as UnitedHealth Group, Kaiser Permanente, Mayo Clinic, and CVS Health can adapt the framework for patient risk stratification, resource allocation, appointment optimization, and clinical decision support. Explainable recommendations improve clinician confidence and facilitate compliance with healthcare governance requirements.In manufacturing and supply chain management, companies such as Tesla, General Motors, Caterpillar, Boeing, and Procter & Gamble can deploy the framework to optimize inventory planning, demand forecasting, supplier risk assessment, predictive maintenance, and logistics optimization. Multi-agent collaboration enables rapid adaptation to disruptions while improving decision consistency across complex supply networks.

The framework is also well suited for telecommunications, insurance, energy, hospitality, and government services, where autonomous AI agents can continuously monitor operational data, generate predictive insights, explain decision logic, and recommend strategic actions. Because the architecture combines high predictive accuracy with transparency, scalability, and collaborative reasoning, it aligns with growing U.S. industry requirements for trustworthy, explainable, and human-centered AI systems.

Overall, the experimental evaluation demonstrates that the proposed Agentic AI framework not only surpasses conventional machine learning models in predictive performance but also delivers explainable, actionable, and business-oriented intelligence. These characteristics make it a strong candidate for deployment in real-world U.S. organizations seeking to leverage advanced AI for data-driven decision-making while maintaining transparency, accountability, and operational resilience.

Conclusion

This study presented a novel **Agentic Artificial Intelligence framework for intelligent business decision-making** by integrating autonomous multi-agent systems, Large Language Models (LLMs), Explainable Artificial Intelligence (XAI), and Business Intelligence (BI) into a unified decision-support architecture. Unlike conventional business analytics systems that primarily focus on predictive accuracy, our proposed framework combines collaborative AI agents with advanced machine learning and explainability techniques to deliver accurate, transparent, and actionable business recommendations. By assigning specialized responsibilities to autonomous agents for data preprocessing, feature engineering, predictive modeling, explanation generation, and business reasoning, the framework enables coordinated intelligence that closely resembles organizational decision-making processes.

Using the publicly available **Online Shoppers Purchasing Intention Dataset** from the UCI Machine Learning Repository, we demonstrated that the proposed framework achieves superior predictive performance while maintaining high interpretability and decision consistency. Among the evaluated machine learning algorithms, **LightGBM** emerged as the best-performing predictive model, and its integration with SHAP, LIME, and LLM-based reasoning further enhanced the quality of business insights. The experimental results indicate that the multi-agent collaboration mechanism not only improves prediction accuracy but also strengthens the transparency, reliability, and business actionability of AI-generated recommendations. These capabilities are particularly important in domains where organizational decisions require both analytical precision and explainable reasoning.

The proposed framework contributes to the existing literature by bridging several important gaps. First, it extends traditional Business Intelligence systems beyond descriptive and predictive analytics by incorporating autonomous multi-agent collaboration capable of adaptive reasoning and coordinated decision-making. Second, it combines explainable AI with LLM-driven contextual interpretation, enabling business users to understand not only **what** prediction has been made but also **why** the prediction was generated and **how** it can be translated into strategic business actions. Third, the framework provides a scalable architecture that integrates structured predictive analytics with natural language reasoning, making advanced AI technologies more accessible to executives, analysts, and organizational decision-makers.

, POLICY-ALIGNED FRAMEWORK FOR HIGH-RISK ENTERPRISE REASONING. *Journal of Engineering Education and Practice*, *2*(1), 48-69. From an industrial perspective, the proposed Agentic AI framework has broad applicability across multiple sectors of the U.S. economy, including retail, e-commerce, banking, financial services, healthcare, manufacturing, supply chain management, insurance, telecommunications, and logistics. Organizations can deploy the framework to support customer behavior analysis, demand forecasting, fraud detection, inventory optimization, personalized marketing, risk assessment, and strategic planning. The combination of predictive intelligence, explainability, and autonomous collaboration allows enterprises to make faster, more informed, and more trustworthy decisions while reducing operational complexity and improving organizational agility.

Despite these promising findings, this research has several limitations. The experimental evaluation was conducted using a single publicly available benchmark dataset representing customer purchasing behavior, which may not fully capture the diversity and complexity of enterprise data encountered across different industries. In addition, the proposed framework was evaluated within an offline analytical environment rather than a real-time production system. The implementation of Large Language Models was designed to provide contextual reasoning and recommendation generation; however, the framework did not explicitly incorporate external organizational knowledge bases, retrieval-augmented generation (RAG), or continuous learning mechanisms that could further improve decision quality in dynamic business environments.

Future research will focus on extending the proposed framework to support multimodal enterprise data by integrating structured databases with textual documents, financial reports, images, sensor streams, and real-time Internet of Things (IoT) data. We also plan to investigate reinforcement learning-based autonomous planning, retrieval-augmented generation, federated learning for privacy-preserving business analytics, graph neural networks for organizational knowledge representation, and human-in-the-loop decision support to further enhance collaborative intelligence. Additionally, large-scale deployment and validation across real-world U.S. industries will be conducted to evaluate the scalability, robustness, and economic impact of the proposed architecture under operational conditions.

In conclusion, this research demonstrates that the integration of Agentic Artificial Intelligence, Large Language Models, Explainable Artificial Intelligence, and Business Intelligence provides a powerful foundation for next-generation enterprise decision-support systems. The proposed framework moves beyond traditional predictive analytics by enabling autonomous collaboration, transparent reasoning, and actionable business intelligence. As organizations continue to embrace digital transformation and increasingly rely on AI-driven strategies, the proposed architecture offers a scalable, trustworthy, and intelligent solution capable of supporting complex business decisions in rapidly evolving environments. We believe that this work provides a significant step toward the realization of autonomous, explainable, and human-centered business intelligence systems that will play a central role in the future of intelligent enterprises.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

[1] Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., & et al. (2021). *On the opportunities and risks of foundation models*. Stanford Center for Research on Foundation Models. https://arxiv.org/abs/2108.07258

[2] Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I., & Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems, 33*, 1877–1901.

[3] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/10.1145/2939672.2939785

[4] Dua, D., & Graff, C. (2019). *UCI machine learning repository*. University of California, Irvine, School of Information and Computer Sciences. https://archive.ics.uci.edu/ml

[5] Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science, 349*(6245), 255–260. https://doi.org/10.1126/science.aaa8415

[6] Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T. Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems, 30*, 3146–3154.

[7] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444. https://doi.org/10.1038/nature14539

[8] Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems, 30*, 4765–4774.

[9] OpenAI. (2023). *GPT-4 technical report*. arXiv. https://arxiv.org/abs/2303.08774

[10] Park, J. S., O'Brien, J., Cai, C., Morris, M., Liang, P., & Bernstein, M. S. (2023). Generative agents: Interactive simulacra of human behavior. *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, 1–22. https://doi.org/10.1145/3586183.3606763

[11] Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: Unbiased boosting with categorical features. *Advances in Neural Information Processing Systems, 31*, 6638–6648.

[12] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144. https://doi.org/10.1145/2939672.2939778

[13] Russell, S., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.

[14] Sakar, C. O., Polat, S. O., Katircioglu, M., & Kastro, Y. (2019). Real-time prediction of online shoppers' purchasing intention using multilayer perceptron and LSTM recurrent neural networks. *Neural Computing and Applications, 31*(10), 6893–6908. https://doi.org/10.1007/s00521-018-3523-0

[15] Sakar, C. O., Polat, S. O., Katircioglu, M., & Kastro, Y. (2019). *Online shoppers purchasing intention dataset*. UCI Machine Learning Repository. https://archive.ics.uci.edu/dataset/468/online+shoppers+purchasing+intention+dataset

[16] Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M. A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., Rodriguez, A., Joulin, A., Grave, E., & Lample, G. (2023). LLaMA: Open and efficient foundation language models. *arXiv*. https://arxiv.org/abs/2302.13971

[17] Watson, H. J. (2019). *Business intelligence: Strategy and big data analytics*. Pearson.

[18] Wooldridge, M. (2009). *An introduction to multiagent systems* (2nd ed.). John Wiley & Sons.

[19] Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K., & Cao, Y. (2023). ReAct: Synergizing reasoning and acting in language models. *International Conference on Learning Representations (ICLR 2023)*. https://arxiv.org/abs/2210.03629

[20] Mia, M. M., Roy, M. K., YASSAR, I. S., Mottalib, M. Y., Yezdani, S., Nijhum, A. M., ... & Uddin, M. K. (2025). Integrating Blockchain Security and Machine Learning for Fraud Detection in the US Banking System. *Emerging Frontiers Library for The American Journal of Engineering and Technology*, *7*(11), 65-76.

[21] YASSAR, I. S. (2026). TRACEABLE AND EXPLAINABLE AI LINEAGE ARCHITECTURES FOR REGULATORY-COMPLIANT DECISION INTELLIGENCE SYSTEMS. *Computers and Education Letters*, *3*(01), 1-18.

[22] Khatun, P., Umam, S., Razzak, R. B., Shamsuddin, I. B., & Salma, N. (2025). A study on the effectiveness of machine learning models for hepatitis prediction. *Scientific Reports*, *15*(1), 30659.

[23] Umam, S., Razzak, R. B., Munni, M. Y., & Rahman, A. (2025). Exploring the non-linear association of daily cigarette consumption behavior and food security-An application of CMP GAM regression. *PLoS One*, *20*(7), e0328109.
[24] Razzak, R. B., & Umam, S. (2025, November). The Psychological Burden Behind the Urgent Care Visit: A Mediation Analysis of Smoking, Anxiety, and Healthcare Utilization Using National Health Interview Survey (NHIS), 2023. In *APHA 2025 Annual Meeting and Expo*. APHA.
[25] Nguyen, A. T. P., Shak, M. S., & Al-Imran, M. (2024). Advancing early skin cancer detection: A comparative analysis of machine learning algorithms for melanoma diagnosis using dermoscopic images. *International Journal of Medical Science and Public Health Research*, *5*(12), 119-133.
[26] Mahmud, F., Das, A. C., Shak, M. S., Rahman, N., Eva, A. A., Mridha, M. F., & Hossen, M. J. (2025). HybridTabNet-QC: A transformer-based clinical feature fusion framework for heart disease risk prediction. *IEEE Open Journal of the Computer Society*, *7*, 1-13.
[27] Das, A. C., Shak, M. S., Rahman, N., Mahmud, F., Shoaib, H. A., & Hossen, M. J. (2026). Cardiovascular disease prediction using variational recurrent autoencoders with uncertainty estimation. *Scientific Reports*.
[28] Das, A. C., Shak, M. S., Rahman, N., Mahmud, F., Shoaib, H. A., & Hossen, M. J. (2026). Cardiovascular disease prediction using variational recurrent autoencoders with uncertainty estimation. *Scientific Reports*.
[29] Das, A. C., Shak, M. S., Rahman, N., Mahmud, F., Eva, A. A., & Hasan, M. N. (2025, May). Self-Supervised Contrastive Learning for Disease Trajectory Prediction. In *2025 5th International Conference on Pervasive Computing and Social Networking (ICPCSN)* (pp. 732-738). IEEE.
[30] Mozumder, M. A. S., Mahmud, F., Shak, M. S., Sultana, N., Rodrigues, G. N., Al Rafi, M., ... & Bhuiyan, M. S. M. (2024). Optimizing customer segmentation in the banking sector: a comparative analysis of machine learning algorithms. *Journal of Computer Science and Technology Studies*, *6*(4), 01-07.
[31] Rahman, M. H., Das, A. C., Shak, M. S., Uddin, M. K., Alam, M. I., Anjum, N., ... & Alam, M. (2024). Transforming customer retention in fintech industry through predictive analytics and machine learning. *The American Journal of Engineering and Technology*, *6*(10), 150-163.
[32] Bhuiyan, R. J., Akter, S., Uddin, A., Shak, M. S., Islam, M. R., Rishad, S. S. I., ... & Hasan-Or-Rashid, M. (2024). Sentiment analysis of customer feedback in the banking sector: A comparative study of machine learning models. *The American Journal of Engineering and Technology*, *6*(10), 54-66.
[33] Razzak, R. B., & Umam, S. (2025, November). Health Equity in Action: Utilizing PRECEDE-PROCEED Model to Address Gun Violence and associated PTSD in Shaw Community, Saint Louis, Missouri. In *APHA 2025 Annual Meeting and Expo*. APHA.
[34] Khatun, P., Umam, S., Razzak, R. B., Shamsuddin, I. B., & Salma, N. (2025). A study on the effectiveness of machine learning models for hepatitis prediction. *Scientific reports*, *15*(1), 30659.
[35] Umam, S., Razzak, R. B., Munni, M. Y., & Rahman, A. (2025). Exploring the non-linear association of daily cigarette consumption behavior and food security-An application of CMP GAM regression. *Plos one*, *20*(7), e0328109.
[36] Adams, R., Grellner, S., Umam, S., & Shacham, E. (2023, November). Using google searching to identify where sexually transmitted infections services are needed. In *APHA 2023 Annual Meeting and Expo*. APHA.
[37] Umam, S., Adams, R., Shacham, E., & Charles, D. L. (2024, October). Predictors of weapon carrying for high school students. In *APHA 2024 Annual Meeting and Expo*. APHA.
[38] YASSAR, I. S. (2023). SCALABLE SDN-BASED ARCHITECTURE FOR LARGE-SCALE ENTERPRISE NETWORK MANAGEMENT. *Insights Sustainable Engineering Practices*, *1*(01), 115-130.