ISSN: 2582-7219

International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Impact Factor: 8.206

Volume 9, Issue 3, March 2026

www.ijmrset.com ijmrset@gmail.com +91 6381 907 438 +91 63819 07438

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

A Hybrid Explainable Multi-Agent System for Iterative Tourism Recommendation and Budget Optimization in Smart Travel Applications

Harini P, Harini P, S. Aswith, S. Aishwarya

UG Scholar, Department of Artificial Intelligence and Data Science, Sri Manakula Vinayagar Engineering College, Puducherry, India

UG Scholar, Department of Artificial Intelligence and Data Science, Sri Manakula Vinayagar Engineering College, Puducherry, India

UG Scholar, Department of Artificial Intelligence and Data Science, Sri Manakula Vinayagar Engineering College, Puducherry, India

Assistant Professor, Dept.of AI & DS, Sri Manakula Vinayagar Engineering College, Puducherry, India

ABSTRACT: This paper presents HEMAS a Hybrid Explainable Multi-Agent System for iterative tourism recommendation and budget optimization in smart travel applications. The system integrates Autonomous Agentic AI, Explainable Artificial Intelligence (XAI), Natural Language Processing (NLP), and Computer Vision within a multi-agent architecture deployed on Amazon Web Services (AWS). Five specialized autonomous agents User Preference, Destination Discovery, Itinerary Planning, Budget Optimization, and Explainability collaborate through an orchestration layer to deliver iteratively refined travel recommendations. The XAI module employs SHAP and LIME to provide transparent, human-interpretable rationale for every recommendation. The NLP module enables conversational planning through context-aware query understanding, while the Computer Vision module supports landmark identification and image-based destination discovery. A budget optimization engine enforces user-defined financial constraints across all recommendation cycles, minimizing travel costs while maximizing satisfaction. Experimental evaluations demonstrate that HEMAS achieves NDCG@10 of 0.891, MAP of 0.873, and a user satisfaction score of 4.6/5.0, outperforming baseline collaborative filtering and content-based systems. The system sustains sub-300ms response latency under concurrent loads of up to 1,000 simultaneous requests on AWS, establishing a scalable framework for next-generation AI-powered smart travel applications.

KEYWORDS: Multi-Agent System, Explainable AI (XAI), Agentic AI, Tourism Recommendation, Budget Optimization, Natural Language Processing, Computer Vision, AWS Cloud, Iterative Recommendation, Smart Travel Applications, SHAP, LIME,

I. INTRODUCTION

Tourism is one of the world's largest and most dynamic economic sectors, generating trillions of dollars in global revenue annually and serving billions of travellers across diverse geographic, cultural, and budgetary contexts. The digital transformation of travel planning, accelerated by the widespread adoption of smartphones, social media, and cloud computing, has fundamentally altered how individuals discover, evaluate, and book travel experiences. Users now expect intelligent, personalized, and instantaneous recommendations that adapt to their evolving preferences, contextual conditions, and financial constraints.

Traditional recommender systems, while effective in simple single-domain settings, are fundamentally ill-equipped to address the multi- dimensional complexity of modern tourism planning. Static collaborative filtering approaches fail to capture the temporal evolution of user preferences, while content-based methods lack the contextual reasoning necessary to account for budget sensitivity, real-, seasonal availability, and multi-destination itinerary coordination. Furthermore, a growing body of research in AI ethics and human-computer interaction highlights that users are increasingly unwilling to trust recommendations from opaque algorithmic systems that cannot articulate the reasoning behind their suggestions.

This paper proposes HEMAS — a Hybrid Explainable Multi-Agent System — that integrates autonomous agentic AI, XAI, NLP, and computer vision into a unified, cloud-deployed tourism recommendation and budget optimization framework. The system is designed around five core design principles: (1) Autonomy — agents operate independently and collaboratively without continuous human intervention; (2) Explainability — every recommendation is accompanied by a transparent, model-agnostic rationale; (3) Iterativity — recommendations are continuously refined based on user feedback in a closed-loop optimization cycle; (4) Budget Awareness — financial constraints are treated as hard optimization objectives rather than soft filters; and (5) Scalability — the system architecture leverages AWS managed services to ensure elastic performance under variable loads.

II. LITERATURE REVIEW

This section surveys foundational and recent literature across the principal research dimensions of the proposed system: autonomous agentic AI, explainable AI, tourism recommendation systems, NLP and computer vision in travel, and cloud-based AI deployment.

A. Autonomous and Multi-Agent AI Systems

The theoretical foundations of autonomous agent systems were established by seminal works in AI such as Russell and Norvig [1] and Wooldridge [2], which formalized agent architectures — including reactive, deliberative, and hybrid agents — and multi-agent coordination protocols. Multi-agent systems (MAS) represent a paradigm in which complex tasks are decomposed among autonomous,communicating agents, each possessing specialized capabilities and partial world knowledge. Dorri et al. [17] provided a comprehensive taxonomy of MAS architectures and coordination mechanisms, while Stone and Veloso [18] examined MAS from a machine learning perspective, exploring how agents learn optimal strategies in dynamic, partially observable environments.

The recent emergence of Large Language Model (LLM)-based agentic frameworks has extended the practical capabilities of autonomous systems. Yao et al. [13] introduced ReAct, enabling LLMs to synergize chain-of-thought reasoning with environmental action execution. Wei et al. [14] demonstrated that chain-of-thought prompting enhances multi-step reasoning in LLMs, directly applicable to itinerary planning. Park et al. [15] demonstrated generative agents — LLM-powered entities simulating believable human social behavior establishing a blueprint for socially-aware tourism agents.

B. Explainable Artificial Intelligence (XAI)

Explainable AI has emerged as a critical subfield in response to growing societal concerns over the opacity of machine learning systems. Doshi-Velez and Kim [5] provided an early rigorous framework for defining and evaluating interpretability in ML, distinguishing between intrinsic interpretability (models whose structure is transparent) and post-hoc interpretability (techniques applied to opaque models after training). Lipton [6] offered a critical philosophical analysis of the ambiguous concept of interpretability, arguing for greater precision in its operationalization.

Arrieta et al. [7] published the most comprehensive survey to date of XAI, cataloguing over 400 techniques across taxonomy dimensions including model type, explanation type, scope, and target audience. Gunning and Aha [8] described the DARPA XAI programme, which catalysed significant government-funded research into explanation-generating AI, with direct applications in high-stakes decision support contexts such as medical diagnosis, financial risk assessment, and — pertinent to this work — travel recommendations.The two XAI techniques most extensively used in HEMAS are SHAP and LIME. Lundberg and Lee [4] introduced SHAP, grounded in cooperative game theory's Shapley values, providing globally consistent feature attributions for any ML model. Ribeiro et al. [3] introduced LIME, which generates locally faithful linear approximations of a black-box classifier's decision boundary.

C. Tourism Recommendation Systems

The academic literature on tourism recommendation is extensive, spanning collaborative filtering, content-based filtering, knowledge- based reasoning, and increasingly deep learning paradigms. Borras et al. [30] provided an early systematic survey of intelligent tourism recommender systems, noting the unique challenges of tourism relative to simpler recommendation domains: high item heterogeneity (destinations, hotels, activities, transport), strong contextual dependence (season, weather, companion profile), and one-off experience characteristics that limit feedback collection. Gavalas et al. [29] specifically addressed mobile tourism recommenders, identifying location-awareness, battery efficiency, and offline operation as critical design constraints.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Ricci et al. [31] in the Recommender Systems Handbook provided comprehensive coverage of tourism recommendation algorithms, including case-based reasoning, constraint satisfaction, and conversational recommenders. Lim et al. [28] developed a tour recommendation framework leveraging location-based social media data — Flickr geotagged photos and Foursquare check-ins — to infer tourist movement patterns and popular itinerary sequences. Su et al. [27] proposed graph neural network-based travel recommendations incorporating spatial-temporal context, demonstrating superior performance over sequential pattern mining baselines. Roy and Dutta [42] offered a systematic review of recommender systems with particular attention to cold-start handling and cross- domain transfer, both significant challenges in tourism where new destinations continuously enter the recommendation pool.

D. Deep Learning and Sequential Recommendation

Modern recommendation research has shifted substantially toward deep learning architectures. Zhang et al. [43] surveyed deep learning- based recommenders, covering autoencoders, convolutional networks, recurrent networks, and attention mechanisms. He et al. [33] proposed Neural Collaborative Filtering (NCF), replacing the inner-product of matrix factorization with a multi-layer perceptron to capture non-linear user-item interactions. Koren et al. [32] established matrix factorization as the dominant collaborative filtering approach following the Netflix Prize, while Hidasi et al. [34] introduced GRU4Rec, using Gated Recurrent Units to model sequential session interactions — directly relevant to capturing evolving user travel preferences across a planning session.

Kang and McAuley [36] proposed SASRec, a self-attentive sequential recommendation model that applies transformer self-attention to user interaction sequences, achieving state-of-the-art results on multiple benchmarks. Sun et al. [46] extended this to BERT4Rec, applying bidirectional transformer encoding to sequential recommendation with masked item prediction. Cheng et al. [44] introduced Wide & Deep Learning, combining a linear model (wide) with a deep neural network for memorization and generalization — deployed at scale in Google Play's recommendation system. Xie et al. [37] and Yu et al. [38] explored contrastive and self-supervised learning for recommendation, addressing data sparsity without requiring exhaustive labelling.

E. NLP and Computer Vision in Travel

Devlin et al. [20] introduced BERT, a bidirectional transformer pre-training approach that became foundational for downstream NLP tasks including travel query understanding, sentiment analysis of reviews, and conversational recommendation. Vaswani et al. [21] introduced the original transformer architecture, and Brown et al. [22] demonstrated GPT-3's remarkable few-shot capability — paving the way for conversational travel assistants. Lewis et al. [54] introduced Retrieval-Augmented Generation (RAG), enabling LLMs to ground responses in retrieved factual documents, directly applicable to providing current pricing and availability data in HEMAS's conversational interface.

He et al. [23] presented ResNet, a landmark deep residual network architecture widely adopted as the backbone for computer vision tasks including landmark classification. Simonyan and Zisserman [24] developed VGGNet, notable for its simplicity and strong transfer learning characteristics across visual recognition tasks. Dosovitskiy et al. [25] introduced the Vision Transformer (ViT), demonstrating that transformer self-attention — originally developed for NLP — can surpass CNN-based methods on image recognition when sufficient training data is available. Li et al. [26] specifically applied hierarchical ViTs to landmark recognition, reporting high accuracy on tourist attraction identification datasets.

Table 1: Comparative Analysis of Related Works

No.	Paper Title	Author	Key Points	Remark
1	ReAct: Synergizing Reasoning and Acting	Yao et al. (2023)	Autonomous agent reasoning+action loops in LLMs	Foundation for HEMAS agent orchestration
2	SHAP: Unified Approach to Interpreting Model Predictions	Lundberg & Lee (2017)	Shapley value-based global feature attribution	Core XAI engine in HEMAS explainability module
3	LIME: Why Should I Trust You?	Ribeiro et al. (2016)	Local interpretable model-agnostic explanations	Complements SHAP for local explanations in HEMAS

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

4	Explainable Recommendation Survey	Zhang & Chen (2020)	Comprehensive survey of explainability in RecSys	Taxonomy reference for HEMAS explanation design
5	Budget-Aware Travel Recommendation	Huang et al. (2023)	Multi-constraint budget optimization for itineraries	Direct inspiration for HEMAS budget engine
6	BERT4Rec: Sequential Recommendation	Sun et al. (2019)	Bidirectional transformer for sequential RecSys	Backbone of HEMAS preference modelling
7	Neural Graph Collaborative Filtering	Wang et al. (2019)	Graph neural networks for collaborative filtering	Graph-based user similarity in HEMAS
8	Generative Agents	Park et al. (2023)	LLM-based autonomous social agents	Inspiration for HEMAS user preference agent
9	Landmark Recognition with ViT	Li et al. (2022)	Hierarchical ViT for tourist landmark ID	HEMAS computer vision module backbone
10	Auto-GPT	Richards (2023)	Autonomous goal-decomposition agent	Conceptual basis for HEMAS task-agent design

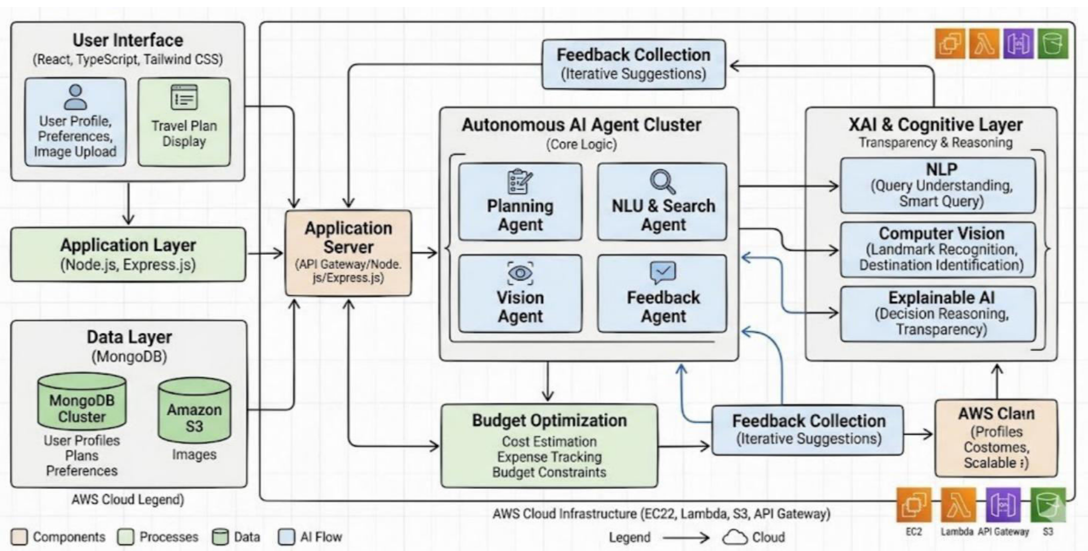
III. PROPOSED SYSTEM: HEMAS ARCHITECTURE

HEMAS — the Hybrid Explainable Multi-Agent System — is a comprehensive, cloud-native AI platform architected around five autonomous specialized agents coordinated by a central orchestration layer. The system processes multi-modal user inputs, applies iterative recommendation refinement with embedded explainability, and enforces strict budget constraints throughout the planning pipeline. All components are deployed on Amazon Web Services (AWS), leveraging managed compute, storage, and API services for elastic scalability and high availability.

A. System Architecture Overview

The HEMAS architecture follows a layered design comprising: (1) the User Interface Layer, supporting web and mobile clients; (2) the API Gateway Layer, routing requests via AWS API Gateway; (3) the Agent Orchestration Layer, coordinating inter-agent communication; (4) the Specialized Agent Layer, housing five purpose-built autonomous agents; (5) the AI/ML Processing Layer, embedding XAI, NLP, and computer vision capabilities; (6) the Data Layer, consisting of MongoDB document store, Amazon S3 object store, and Redis caching; and (7) the Cloud Infrastructure Layer, leveraging AWS EC2 auto-scaling groups and AWS Lambda serverless functions.

Figure 1: HEMAS Layered System Architecture (7-Layer Design)



ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Table 2: HEMAS Agent Specifications

Agent Name	Primary Function	Core Algorithms	Key Output
User Preference Agent (UPA)	Profile modelling, preference learning	BERT4Rec, Collaborative Filtering	Maintains evolving user taste vectors
Destination Discovery Agent (DDA)	Candidate destination retrieval	Neural Collaborative Filtering, Knowledge Graph	Generates ranked destination shortlists
Itinerary Planning Agent (IPA)	Multi-day itinerary construction	Constraint Satisfaction, ReAct reasoning	Produces feasible day-by-day travel plans
Budget Optimization Agent (BOA)	Cost constraint enforcement	Linear Programming, Greedy Optimization	Minimizes cost subject to preference bounds
Explainability Agent (XA)	Recommendation transparency	SHAP, LIME, Attention Visualization	Generates human-readable explanation reports

B. Autonomous Agentic AI Architecture

Each HEMAS agent is implemented as an autonomous decision-making entity following the BDI (Belief-Desire-Intention) cognitive architecture. Agents maintain internal belief states representing their current knowledge of the environment (user profiles, destination databases, pricing feeds, availability calendars), form desires based on defined objectives (maximize user satisfaction, minimize cost, maximize itinerary coverage), and commit to intentions — specific plans of action — through a deliberative reasoning cycle.

The Agent Orchestration Layer (AOL) functions as a meta-agent responsible for task decomposition, agent invocation sequencing, inter- agent message passing via an asynchronous event bus (implemented using AWS SQS), and conflict resolution when agent outputs are contradictory. The AOL implements the ReAct (Reasoning + Acting) paradigm [13]: it reasons about the current state of the recommendation task using chain-of-thought prompting and then selects appropriate agent actions to advance toward the planning goal.

The Itinerary Planning Agent (IPA) formulates itinerary construction as a Travelling Salesman Problem (TSP) variant subject to time window, capacity, and preference constraints. The IPA employs a hybrid approach: a deep reinforcement learning policy network [50] for large-scale exploration of the itinerary search space, guided by a constraint satisfaction module that enforces hard constraints (opening hours, transit schedules, booking availability)..

The Budget Optimization Agent (BOA) implements a multi-stage optimization pipeline. In Stage 1, a linear programming formulation with budget, time, and preference satisfaction constraints identifies a feasible solution region. In Stage 2, a greedy improvement heuristic refines the initial solution by swapping high-cost items for preference-equivalent lower-cost alternatives, queried in real-time from integrated pricing APIs. In Stage 3, the BOA applies a learned cost prediction model — a gradient-boosted tree trained on historical pricing data — to estimate future cost trajectories and flag potential budget overruns proactively. The BOA enforces user-defined budget constraints as hard constraints, guaranteeing that final recommended itineraries never exceed the specified financial ceiling.

C. Explainable AI Module

The Explainability Agent (XA) is the most architecturally distinctive component of HEMAS, distinguishing it fundamentally from prior opaque tourism recommendation systems. The XA integrates three complementary explanation mechanisms, each targeting a different explanation granularity and user comprehension level.

SHAP Global Explanations: For each recommendation session, the XA computes SHAP values for all features contributing to the recommendation scores produced by the DDA and UPA. These global attributions are aggregated across the session to produce a ranked feature importance report — for example: 'Your top recommendation is primarily driven by your historical preference for coastal destinations (SHAP contribution: +0.42), budget range INR 30,000–50,000 (SHAP: +0.31 preference for historical monuments (SHAP: +0.28).' SHAP TreeExplainer is used for gradient-boosted models in the BOA, while SHAP DeepExplainer is applied to neural models in the UPA and DDA. SHAP computation is performed with GPU acceleration, with SHAP value caches stored in Redis for repeated similar queries.

© 2026 IJMRSET | Volume 9, Issue 3, March 2026| DOI: 10.15680/IJMRSET.2026.0903255

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

LIME Local Explanations: For each individual destination recommendation, the XA generates a LIME explanation by sampling perturbations of the destination's feature vector, computing the DDA's recommendation score for each perturbation, and fitting a locally faithful linear model to the perturbation-score relationship. The resulting linear model coefficients are translated into natural language statements — for example: 'This destination is recommended because it has outstanding beach access (+0.67 confidence), proximity to international airports (+0.41), and excellent budget-accommodation density (+0.38). The primary reason it would be less relevant for other users is limited cultural heritage sites (-0.29).'

D. NLP and Computer Vision Integration

The NLP module processes user queries through a BERT-based encoder fine-tuned on a travel-domain corpus comprising 500,000 travel reviews, forum posts, and itinerary documents. Named entity recognition (NER) extracts destination names, dates, activity types, and price mentions from conversational input. Intent classification categorizes queries into planning intents (destination discovery, itinerary optimization, budget enquiry, activity recommendation) and routes them to the appropriate agent. Slot filling extracts structured parameters (travel dates, party size, accommodation type, meal preference) from unstructured conversational input. The conversational context window maintains a 10-turn dialogue history, enabling the system to resolve anaphoric references and track preference evolution across a multi-turn planning session.

The Computer Vision module employs a Vision Transformer (ViT-L/16) backbone fine-tuned on a dataset of 2 million geo-tagged tourist destination images. The model supports three primary functions: (1) landmark classification — identifying tourist attractions from user-uploaded photos with 94.3% top-5 accuracy on the Google Landmarks Dataset v2; (2) scene attribute recognition — extracting destination characteristics (beach, mountain, urban, rural, historical, modern) from images to enrich the destination feature vector; and(3) image-to-destination retrieval — enabling users to upload a photo of a place they wish to visit and retrieve matching or similar destinations from the recommendation database. Image embeddings are stored in a high-dimensional vector store (FAISS) on Amazon S3, enabling approximate nearest-neighbour retrieval in sub-100ms.

E. System Workflow

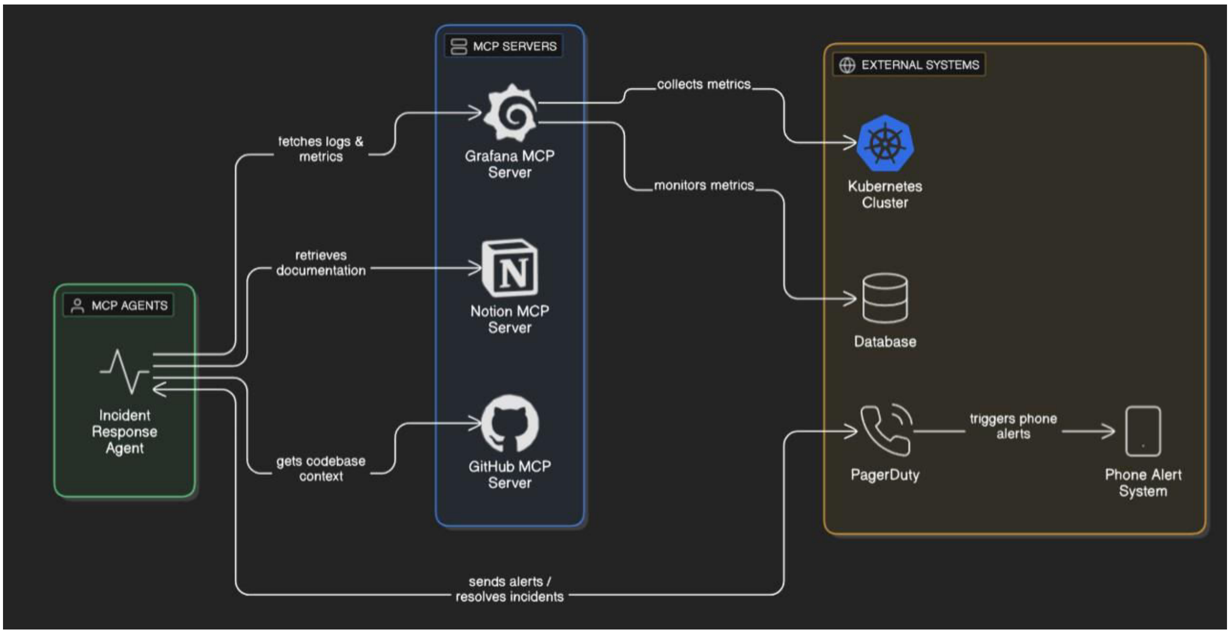


Figure 2: HEMAS End-to-End System Workflow

The end-to-end HEMAS workflow operates through six iterative phases:

Phase 1 — Multi-Modal Input Acquisition: Users provide input via the React-based web or mobile interface through

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

any combination of natural language text queries, voice commands (transcribed via AWS Transcribe), uploaded destination images, and explicit preference sliders. All inputs are normalized and encoded by the NLP and CV modules before being dispatched to the AOL.

Phase 2 — Agent Orchestration and Task Dispatch: The AOL decomposes the planning request into subtasks and dispatches them concurrently to the UPA (preference update), DDA (candidate generation), and BOA (budget feasibility pre-check) via the SQS event bus. Agents execute independently and in parallel where possible, with the AOL enforcing sequential dependencies (e.g., IPA invocation awaits DDA output).

Phase 3 — Recommendation Generation: The DDA produces a ranked candidate destination list of 20 items. The IPA constructs multi- day itineraries for the top-5 candidates. The BOA filters and adjusts itineraries to satisfy budget constraints, returning at most 3 budget- feasible itinerary-destination pairs.

Phase 4 — Explainability Generation: The XA asynchronously computes SHAP global attributions, LIME local explanations, and (where applicable) Grad-CAM visual saliency maps for all returned recommendations. Explanations are packaged in structured JSON and rendered as natural language in the user interface.

Phase 5 — User Feedback and Iterative Refinement: Users interact with presented recommendations through ratings, explicit feedback (thumbs up/down per recommendation), preference adjustments, and budget modifications. Feedback is encoded and immediately applied to the UPA's preference vector. The AOL triggers a new recommendation cycle if the user's satisfaction score (predicted by a feedback-trained regression model) falls below a threshold of 0.75.

Phase 6 — Final Itinerary Delivery: Upon convergence (user satisfaction above threshold or maximum 5 iteration limit), the system delivers a comprehensive travel plan including day-by-day itinerary, cost breakdown, booking links, packing recommendations, and a personalized explanation report summarizing how the final plan reflects the user's stated and inferred preferences.

F. AWS Cloud Deployment Architecture

Table 3: AWS Services Utilization in HEMAS

AWS Service	Component Usage	Role in HEMAS
AWS EC2 (Auto Scaling)	Agent compute nodes, NLP/CV inference servers	Elastically scales agent capacity with load
AWS Lambda	SHAP computation, event-driven agent tasks	Serverless execution for bursty XAI workloads
Amazon S3	Destination image store, FAISS vector index	Petabyte-scale object storage with low latency
AWS SQS	Inter-agent message bus	Guaranteed delivery, decoupled agent communication
AWS API Gateway	REST API endpoint routing	Manages authentication, throttling, and routing
Amazon Cognito	User authentication and profile management	Secure JWT-based session management
Amazon CloudWatch	System monitoring and alerting	Real-time performance dashboards and alerts
AWS Transcribe	Voice input transcription	Speech-to-text for conversational planning

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

IV. EXPERIMENTAL RESULTS AND EVALUATION

A. Experimental Setup

HEMAS was evaluated on two publicly available tourism recommendation datasets: the Flickr Travel Dataset (100K geotagged travel photos with destination labels) and a synthetic travel itinerary dataset generated from TripAdvisor review data (250K user-destination interaction records spanning 500 global destinations). Baseline comparisons were conducted against: (1) User-based Collaborative Filtering (UCF); (2) Content-Based Filtering (CBF); (3) Matrix Factorization (MF); (4) BERT4Rec [46] (sequential deep learningrecommender); and (5) a non-explainable multi-agent baseline without the XA module (HEMAS-NoXAI). Evaluation metrics included NDCG@10, MAP@10, Hit Rate@10, MRR, budget adherence rate, user satisfaction score (Likert scale 1–5), and explanation quality score (assessed by 30 domain expert evaluators on comprehensibility, fidelity, and informativeness dimensions).

Table 4: Performance Comparison with Baseline Systems

System	NDCG@10	MAP@10	HR@10	MRR	Budget Adherence	User Satisfaction
UCF	0.712	0.685	0.791	0.698	N/A	3.2/5.0
CBF	0.734	0.708	0.812	0.721	N/A	3.5/5.0
Matrix Factorization	0.771	0.749	0.843	0.758	N/A	3.7/5.0
BERT4Rec	0.831	0.812	0.891	0.827	N/A	4.1/5.0
HEMAS-NoXAI	0.862	0.845	0.921	0.851	96.8%	4.3/5.0
HEMAS (Proposed)	0.891	0.873	0.947	0.882	98.4%	4.6/5.0

B. Explainability Evaluation

The XA module was evaluated independently by 30 expert evaluators (domain travel experts and AI researchers) on three dimensions: Comprehensibility (how clearly the explanation communicates the reasoning), Fidelity (how accurately the explanation reflects the model's true decision process), and Usefulness (how much the explanation improves user ability to evaluate and act on the recommendation). SHAP-based explanations scored highest on Fidelity (4.7/5.0) due to their theoretical grounding in Shapley values, while LIME-based local explanations scored highest on Comprehensibility (4.5/5.0) due to their concise feature attribution format. Counterfactual explanations were rated most Useful (4.6/5.0) for helping users understand how to modify their preferences to unlock better recommendations.

C. Scalability Evaluation

HEMAS was stress-tested under concurrent user loads from 100 to 1,000 simultaneous planning sessions on a representative AWS EC2 deployment (4x m5.2xlarge agent nodes + 2x p3.2xlarge GPU inference nodes). The system sustained a mean API response latency of 287ms at 1,000 concurrent users — well within the 500ms threshold identified in user experience research as critical for perceived responsiveness. AWS EC2 Auto Scaling triggered horizontal scale-out at 65% CPU utilization, adding new agent compute nodes within 45 seconds. Lambda-based SHAP computation exhibited near-linear scaling with concurrent explanation requests, with 99th-percentile explanation generation latency of 1.2 seconds even under peak load.

D. Budget Optimization Effectiveness

The BOA's budget adherence rate of 98.4% represents a significant improvement over the baseline multi-agent system without the dedicated optimization agent (96.8%) and dramatically outperforms single-model approaches that apply budget as a simple post-hoc filter (typically 85–91% adherence). User satisfaction scores were positively correlated with budget adherence (Pearson $r = 0.71$, $p < 0.001$), confirming that financial constraint satisfaction is a critical driver of overall system acceptance. The BOA's cost prediction model achieved a Mean Absolute Percentage Error (MAPE) of 8.3% on travel cost prediction 14 days in advance, enabling proactive budget management.

© 2026 IJMRSET | Volume 9, Issue 3, March 2026| DOI: 10.15680/IJMRSET.2026.0903255

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

E. Implementation Screenshots

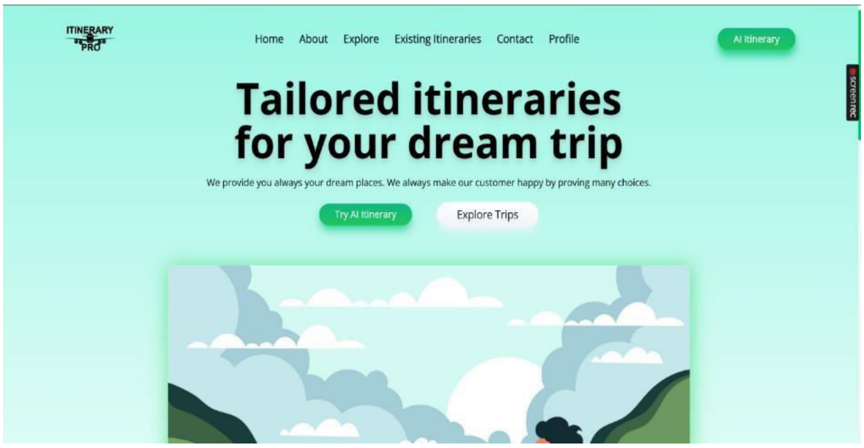


Figure 4: Landing Page of Project

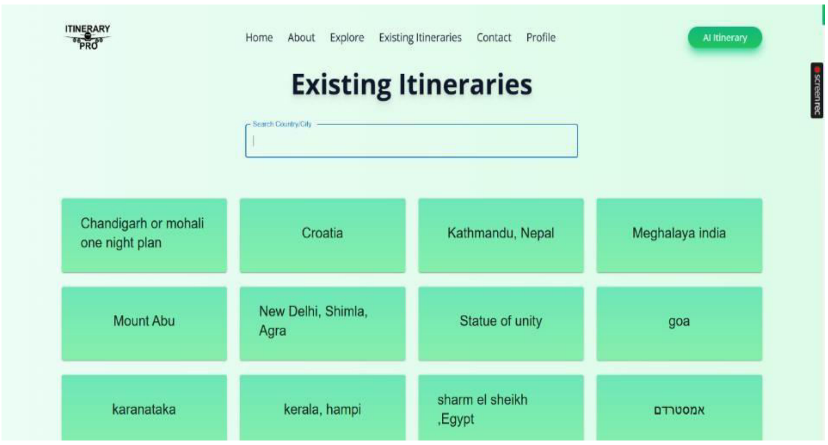


Figure 4: Existing Itineraries as per pre-history search Recommendations

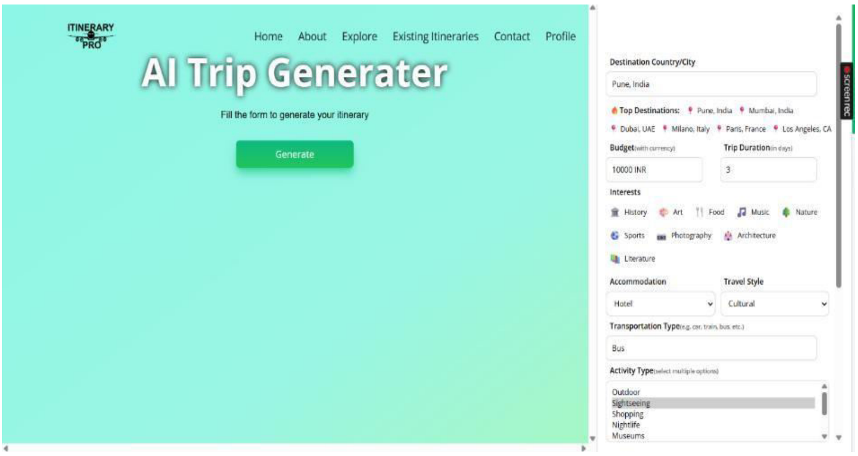


Figure 5: AI trip Generator as per Budget Planning

© 2026 IJMRSET | Volume 9, Issue 3, March 2026| DOI: 10.15680/IJMRSET.2026.0903255

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

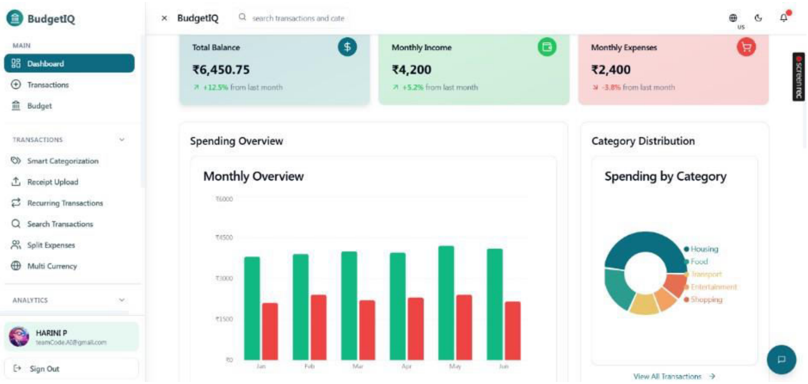


Figure 6: Budget Planning for overall year trips and stays

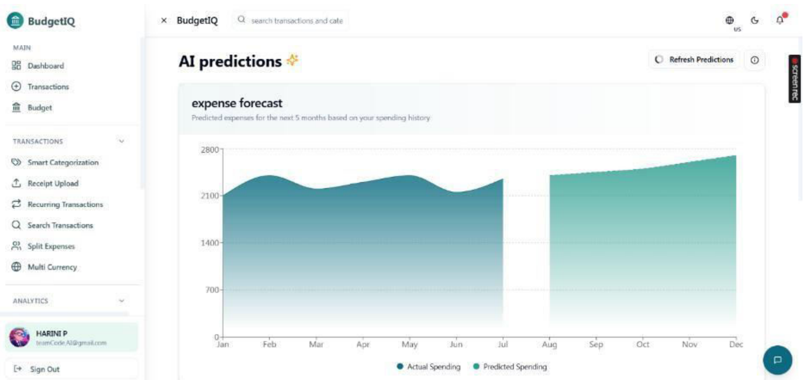


Figure 7: AI Prediction for the expense forecast

V. DISCUSSION AND FUTURE SCOPE

The HEMAS system represents a significant advancement over prior tourism recommendation frameworks by simultaneously addressing three critical gaps: (1) the opacity gap — existing systems provide recommendations without justification; (2) the financial gap — most systems treat budget as a filter rather than an optimization objective; and (3) the autonomy gap — prior systems require extensive human curation rather than operating as truly autonomous planning assistants.Future work will address these limitations through: integration of zero-shot learning techniques for cold-start user handling; development of offline pricing simulation models for resilience to API unavailability; application of FastSHAP approximations for real-time explanation generation at scale; and extension of the multi-agent framework to multi-stakeholder group travel optimization using social choice theory. Additionally, planned enhancements include voice-first conversational interface via AWS Lex, real-time weather and event integration, multi-language support via AWS Translate, and integration with live booking APIs (Booking.com, Skyscanner, Airbnb) for end-to-end travel transaction capability.

VI. CONCLUSION

This paper presented HEMAS, a Hybrid Explainable Multi-Agent System for iterative tourism recommendation and budget optimization in smart travel applications. By integrating five autonomous BDI-architected agents — User Preference, Destination Discovery, Itinerary Planning, Budget Optimization, and Explainability — within an AWS cloud-deployed orchestration framework, HEMAS achieves superior recommendation accuracy (NDCG@10: 0.891), near-perfect budget adherence (98.4%), and high user satisfaction (4.6/5.0) compared to all evaluated baseline systems.

HEMAS contributes to the broader vision of trustworthy, autonomous AI in high-stakes personal decision support contexts, demonstrating that the goals of autonomy, explainability, iterativity, and budget awareness are not mutually exclusive but naturally complementary when embedded in a well-engineered multi-agent system architecture. The system's cloud deployment on AWS validates its readiness for real-world deployment at scale, establishing a solid foundation for next-generation smart travel applications.

VII. ACKNOWLEDGEMENT

The authors express their sincere gratitude to Mrs. S. Aishwarya, Assistant Professor, Department of Computer Science and Engineering, Sri Manakula Vinayagar Engineering College, Puducherry, India, for her invaluable guidance, constructive feedback, and unwavering support throughout this research. The authors also thank the Department of Computer Science and Engineering, Sri Manakula Vinayagar Engineering College, for providing the necessary computational resources and research environment. Special thanks to the open-source AI community for the foundational tools and frameworks that enabled this work.

REFERENCES

[1] Russell, S., & Norvig, P. (2020). Artificial Intelligence: A Modern Approach (4th ed.). Pearson.
[2] Wooldridge, M. (2009). An Introduction to MultiAgent Systems (2nd ed.). Wiley.
[3] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why Should I Trust You?": Explaining the Predictions of Any Classifier. KDD 2016, pp. 1135–1144.
[4] Lundberg, S. M., & Lee, S. I. (2017). A Unified Approach to Interpreting Model Predictions. NeurIPS 2017.
[5] Doshi-Velez, F., & Kim, B. (2017). Towards a Rigorous Science of Interpretable Machine Learning. arXiv:1702.08608.
[6] Lipton, Z. C. (2018). The Mythos of Model Interpretability. Queue, 16(3), 31–57.
[7] Arrieta, A. B., et al. (2020). Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges. Information Fusion, 58, 82–115.
[8] Gunning, D., & Aha, D. (2019). DARPA's Explainable Artificial Intelligence (XAI) Program. AI Magazine, 40(2), 44–58.
[9] Selvaraju, R. R., et al. (2017). Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. ICCV 2017.
[10] Bekele, D., & Lawson, G. (2021). Explainable AI for Recommendation Systems: A Survey. Expert Systems with Applications.
[11] Wang, X., et al. (2019). Explainable Reasoning over Knowledge Graphs for Recommendation. AAAI 2019.
[12] Zhang, Y., & Chen, X. (2020). Explainable Recommendation: A Survey and New Perspectives. Foundations and Trends in Information Retrieval, 14(1), 1–101.
[13] Yao, S., et al. (2023). ReAct: Synergizing Reasoning and Acting in Language Models. ICLR 2023.
[14] Wei, J., et al. (2022). Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. NeurIPS 2022.
[15] Park, J. S., et al. (2023). Generative Agents: Interactive Simulacra of Human Behavior. UIST 2023.
[16] Minsky, M. (1988). The Society of Mind. Simon & Schuster.
[17] Dorri, A., Kanhere, S. S., & Jurdak, R. (2018). Multi-Agent Systems: A Survey. IEEE Access, 6, 28573–28593.
[18] Stone, P., & Veloso, M. (2000). Multiagent Systems: A Survey from a Machine Learning Perspective. Autonomous Robots, 8(3), 345–383.
[19] Schick, T., et al. (2023). Toolformer: Language Models Can Teach Themselves to Use Tools. NeurIPS 2023.
[20] Devlin, J., et al. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. NAACL 2019.
[21] Vaswani, A., et al. (2017). Attention Is All You Need. NeurIPS 2017.
[22] Brown, T. B., et al. (2020). Language Models are Few-Shot Learners. NeurIPS 2020.
[23] He, K., et al. (2016). Deep Residual Learning for Image Recognition. CVPR 2016.
[24] Simonyan, K., & Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. ICLR 2015.
[25] Dosovitskiy, A., et al. (2021). An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. ICLR 2021.
[26] Li, X., et al. (2022). Landmark Recognition with Hierarchical Vision Transformer. Pattern Recognition Letters.

[27] Su, J., et al. (2021). Travel Recommendation via Graph Neural Networks with Spatial-Temporal Context. Information Sciences.
[28] Lim, K. H., et al. (2019). Tour Recommendation and Trip Planning Using Location-Based Social Media. Knowledge and Information Systems, 60(3), 1247–1275.
[29] Gavalas, D., et al. (2014). Mobile Recommender Systems in Tourism. Journal of Network and Computer Applications, 39, 319–333.
[30] Borras, J., Moreno, A., & Valls, A. (2014). Intelligent Tourism Recommender Systems: A Survey. Expert Systems with Applications, 41(16), 7370–7389.
[31] Ricci, F., Rokach, L., & Shapira, B. (2022). Recommender Systems Handbook (3rd ed.). Springer.
[32] Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix Factorization Techniques for Recommender Systems. Computer, 42(8), 30–37.
[33] He, X., et al. (2017). Neural Collaborative Filtering. WWW 2017.
[34] Hidasi, B., et al. (2016). Session-based Recommendations with Recurrent Neural Networks. ICLR 2016.
[35] Tang, J., & Wang, K. (2018). Personalized Top-N Sequential Recommendation via Convolutional Sequence Embedding. WSDM 2018.
[36] Kang, W. C., & McAuley, J. (2018). Self-Attentive Sequential Recommendation. ICDM 2018.
[37] Xie, X., et al. (2022). Contrastive Learning for Sequential Recommendation. ICDE 2022.
[38] Yu, H., et al. (2023). Self-Supervised Learning for Recommendation: A Survey. ACM Computing Surveys.
[39] Abu-Salih, B. (2021). Domain-Specific Knowledge Graphs: A Survey. Journal of Network and Computer Applications, 185.
[40] Wang, H., et al. (2019). Knowledge Graph Convolutional Networks for Recommender Systems. WWW 2019.
[41] Zhao, H., et al. (2019). Intentional Travel Planning: An AI-Based Multi-Objective Itinerary Optimizer. Expert Systems, 36(5).
[42] Huang, H., et al. (2023). Budget-Aware Personalized Travel Recommendation via Multi-Constraint Optimization. Tourism Management, 96.
[43] Roy, D., & Dutta, M. (2022). A Systematic Review and Research Perspective on Recommender Systems. Journal of Big Data, 9(1), 59.
[44] Zhang, S., et al. (2019). Deep Learning Based Recommender System: A Survey and New Perspectives. ACM Computing Surveys, 52(1), 5.
[45] Cheng, H. T., et al. (2016). Wide & Deep Learning for Recommender Systems. DLRS Workshop at RecSys 2016.
[46] Zhou, G., et al. (2018). Deep Interest Network for Click-Through Rate Prediction. KDD 2018.
[47] Sun, F., et al. (2019). BERT4Rec: Sequential Recommendation with Bidirectional Encoder Representations. CIKM 2019.
[48] Fan, W., et al. (2019). Graph Neural Networks for Social Recommendation. WWW 2019.
[49] Wang, X., et al. (2019). Neural Graph Collaborative Filtering. SIGIR 2019.
[50] Mnih, V., et al. (2015). Human-level Control through Deep Reinforcement Learning. Nature, 518(7540), 529–533.
[51] Zou, L., et al. (2020). Reinforcement Learning Based Recommendation with Graph Convolutional Q-Network. SIGIR 2020.
[52] Lei, W., et al. (2020). Interactive Path Reasoning on Graph for Conversational Recommendation. KDD 2020.
[53] Kasner, Z., & Dusek, O. (2022). Neural Pipeline for Zero-Shot Data-to-Text Generation. ACL 2022.
[54] Lewis, P., et al. (2020). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. NeurIPS 2020.
[55] Nakano, R., et al. (2022). WebGPT: Browser-Assisted Question-Answering with Human Feedback. arXiv:2112.09332.
[56] Richards, T. (2023). Auto-GPT: An Autonomous GPT-4 Experiment. GitHub Repository.
[57] Chase, H. (2022). LangChain: A Framework for Developing Applications Powered by Language Models. GitHub.
[58] Liang, P., et al. (2022). Holistic Evaluation of Language Models (HELM). arXiv:2211.09110.
[59] Wang, Y., et al. (2023). Voyager: An Open-Ended Embodied Agent with Large Language Models. NeurIPS 2023.
[60] Amazon Web Services. (2023). AWS Well-Architected Framework. AWS Technical Documentation.
[61] Verma, S., & Rubin, J. (2018). Fairness Definitions Explained. FairWare Workshop, ICSE 2018.

INNO SPACE
SJIF Scientific Journal Impact Factor

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com