IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 5, May 2025

DOI: 10.17148/IARJSET.2025.125334

AI - Based Travel Itinerary

Adarsh Sainath H[1], Chinmay Gowda B S[2], Jayanthan B N[3], Kushal Gowda G[4], Syeda Amira Hussaini[5]

Department of CSE, Maharaja Institute Of Technology, Mysuru, Karnataka, Affiliated to Visvesvaraya Technological University, Belagavi, Karnataka[1-5]

Abstract: As international tourism expands and synthetic intelligence era advances, smart travel planning offerings have emerged as a significant research awareness. Within dynamic actual- global journey scenarios with multi-dimensional constraints, offerings that guide users in automatically developing practical and customized journey itineraries must cope with three key objectives: Rationality, Comprehensiveness, and Personalization. However, existing systems with rule-primarily based combos or LLM-primarily based planning strategies conflict to fully fulfill these criteria. To triumph over the demanding situations, we introduce "TravelAgent", a tour making plans system powered with the aid of massive language fashions (LLMs) designed to offer affordable, comprehensive, and customized travel itineraries grounded in dynamic situations. TravelAgent incorporates four modules: Tool-utilization, Recommendation, Planning, and Memory Module. We evaluate TravelAgent's performance with human and simulated customers, demonstrating its typical effectiveness in three criteria and confirming the accuracy of customized tips.

I. INTRODUCTION

With the boom of worldwide tourism, the call for for automatic journey itinerary technology services has surged. Concurrently, tourists are searching for various and personalized studies increasingly. Early platforms (Wikitravellers, 2024; FrommerMedia LLC, 2024) provide static steerage without advice or making plans. Later structures (Roadtrippers, LLC, 2024; Expedia Group, 2024; Booking Holdings Inc., 2024) offer classified recommendation and reserving services however heavily depend on guide choice and rule- primarily based mixtures, unable to autonomously generate itineraries and adapt to dynamic situations.

Recent advancements in artificial intelligence, specially LLMs, have promoted the development of clever journey offerings. However, in spite of their effectiveness in conventional duties, LLMs face challenges in dynamic, actual-international programs including personalized tour planning, which calls for thinking about more than one constraints and real-time records. The TravelPlanner benchmark (Xie et al., 2024) assesses language agents' talents in dealing with those constraints the use of a dataset of journey eventualities. Results exhibit that even cutting-edge

LLMs (OpenAI, 2022; Achiam et al., 2023; Team et al., 2023; Jiang et al., 2023, 2024) and planning strategies (Wei et al., 2022; Yao et al., 2022; Shinn et al., 2024) war to satisfy all constraints and produce viable plans.

Despite rationality concerns, travel offerings designed for real-international eventualities require complete itineraries incorporating dynamic updates, great-grained information, and large contextual facts. Additionally, personalization is also critical for reinforcing the first-class of offerings. In this study, we set up 3 evaluation criteria for itineraries generated via tour making plans: **Rationality, Comprehensiveness, and Personalization**. Each class introduces implementation challenges that journey service systems need to deal with:

1. **Rationality** - How to version constraints in tour eventualities and plan rational itineraries underneath these constraints?
2. **Comprehensiveness** - How to provide complete tour offerings, which includes advice and making plans, make actual-time, nice-grained, and interesting itineraries?
3. **Personalization** - How to discover and leverage implicit consumer personalization statistics to supply personalized pointers and planning services?

In this paper, we propose TravelAgent, a singular tour carrier gadget powered by way of LLMs and an advanced information processing framework to tackle the mentioned core challenges and meet the established assessment criteria. TravelAgent features 4 distinct modules: the **Tool- Usage Module, Recommendation Module, Planning Module, and Memory Module**; each is meticulously designed and collaborates seamlessly with others. We version the tour constraints and devise efficient algorithms to beautify the Rationality of the starting stage.

DOI: 10.17148/IARJSET.2025.125334

Additionally, the Tool-usage Module includes a pool of real-time tools; we address the comprehension problem with records insights within the Memory Module. Finally, we gift a brand new personalized advice framework for LLM-based totally retailers to sell Personalization.

Furthermore, we rigorously evaluate the effectiveness and fine of TravelAgent through two awesome methodologies: **Overall Evaluation with human customers and Personalization Evaluation with simulated users**. Overall Evaluation assesses TravelAgent in phrases of the 3 criteria via precise case research. Personalization Evaluation delves deeper into Personalization, the usage of simulated users to supply streaming behavioral information for predictive duties in appeal tips. Findings imply that TravelAgent continuously offers excessive performance in Rationality, Comprehensiveness, and Personalization across multiple tour eventualities. Our **contributions** encompass:

- **System:** We outline 3 objective standards for tour offerings and create TravelAgent, a singular AI-pushed gadget designed for computerized personalized tour making plans service to cope with the targets.
- **Method:** For Rationality, we model tour constraints and increase a spatiotemporal aware making plans algorithm. For Comprehensiveness, we set up a real- time records mechanism stronger by using more than one equipment. For Personalization, we introduce an agent-primarily based customized recommendation framework with continuous mastering of implicit information.
- **Evaluation:** We conduct human case studies to assess the general performance of TravelAgent, which surpasses the baseline GPT-4+ agent throughout 3 criteria. Further, we undertake behavioral prediction obligations with simulated customers and discover that our personalized advice framework famous lower errors fees than baseline LLM-based advice methods.

II. RELATED WORK

LLM-primarily based Agent: LLM-primarily based agents have demonstrated vast capability in executing multi-step duties (Wang et al., 2024b), showcasing their capability in fields which include training (Wang et al., 2024b; Swan et al., 2023), healthcare (Zhang et al., 2023), and emotional help (Hasan et al., 2023). To decorate their making plans ability, a few researchers suggest project decomposition techniques (Shen et al., 2024; Wang et al., 2023; Singh et al., 2023), at the same time as others recognition on step-by means of- step reasoning policies for dynamic making plans (Wei et al., 2022; Yao et al., 2024; Besta et al., 2024; Chen et al., 2022; Yao et al., 2022). Additionally, thru reminiscence creation (Zhong et al., 2024; Liu et al., 2023b; Hu et al., 2023; Zheng et al., 2023; Huang et al., 2023) and device use (Schick et al., 2024; Qin et al., 2023; Lu et al., 2024; Wang et al., 2024a; Yuan et al., 2024), LLM-based totally sellers can leverage historic and external assets to better variation, making it a promising tool for tour services.

LLM as Recommendation System: With the arrival of LLMs, researchers endeavor to leverage LLMs for direct hints. Two principal strategies have emerged: Direct Recommendation (Dai et al., 2023; Hou et al., 2024; Kang et al., 2023; Wang and Lim, 2023) and Decode and Encode (Zhou et al., 2024; Yang et al., 2023; Liu et al., 2023a). The former prompts LLMs with customers' behavioral histories to suggest immediately, whilst the latter decodes a one-time person profile from the remaining interplay for subsequent guidelines. However, both methods have obstacles in taking pictures long-term consumer preferences thru more than one interaction, resulting in suboptimal effectiveness.

1. The TravelAgent System Design

TravelAgent device is architected into five key modules: **(1) Tool-usage Module**, which integrates diverse utilities drawing from real-time dynamic data resources or arithmetic algorithms; **(2)Recommendation Module**, which grants personalized tour guidelines tailor-made to person possibilities; **(3) Planning Module**, which includes finances and path planning functionalities, leveraging budget plan end result to guide detailed path plans touchy to time and geographic constraints; **(4) Memory Module**, which help to continually examine user possibilities as a comprehensive repository of journey information and person profiles.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⩋ Peer-reviewed & Refereed journal ⩋ Vol. 12, Issue 5, May 2025

DOI: 10.17148/IARJSET.2025.125334

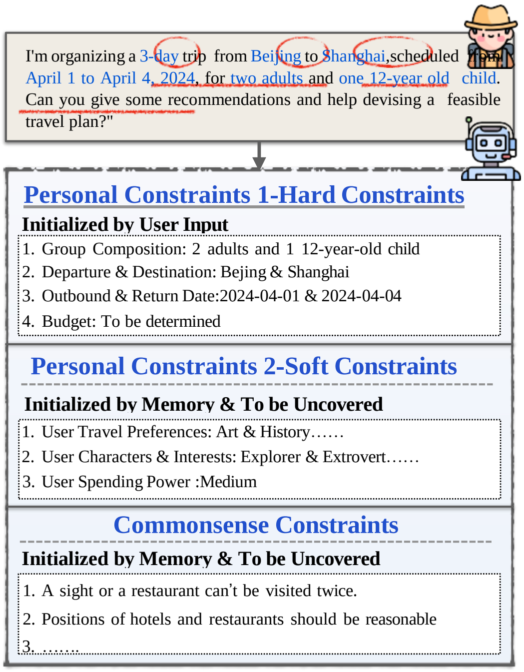

Figure 1: Initial constraints modeling with user input and Memory Module.

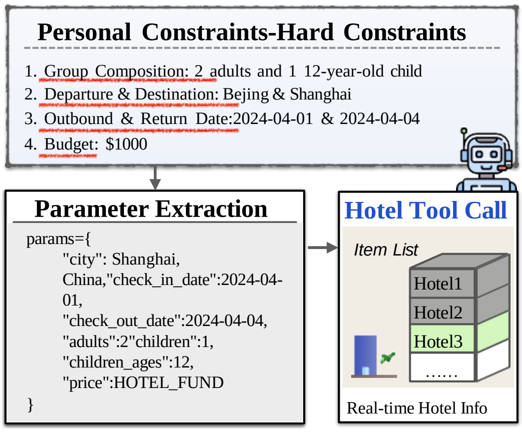

Figure 2: An example of Hotel Tool call. Appendix A provides a comprehensive overview of all types and functions of tools.

1.1. Constrains Modeling

The center assignment in travel services lies in handling multi-dimensional constraints in complicated real-global travel eventualities. In this paintings, we categorize these constraints into - **(1) Personal Constraints**, which might be unique to man or woman customers and consist of Hard Constraints and Soft Constraints; **(2) Commonsense Constraints**, which encompass standard travel expertise applicable across diverse situations and beneficial to all users. Figure 1 illustrates an initial constraints example.

- **Personal Constraints 1 - Hard Constraints** are both person and state of affairs-particular, which should be strictly accompanied. These non-negotiable factors, which include the outbound-/go back date, are without delay furnished by users in the modern-day situation and mirror their travel targets that are crucial for Rationality and Personalization of TravelAgent.
- **Personal Constraints 2 - Soft Constraints** are consumer-precise however move-scenarios, initiated by historical insights from Memory Module. Instead of explicitly provided via users, they may be exposed steadily by TravelAgent throughout a couple of eventualities and evolve with consumer interactions. These personalized elements

DOI: 10.17148/IARJSET.2025.125334

are not obligatory however substantially have an effect on Personalization and user pleasure.

- **Commonsense Constraints** are trendy understanding relevant to customers and eventualities, which might be initiated with the aid of Memory Module and constantly discovered by means of TravelAgent. They are critical for Rationality of itineraries and journey stories, consisting of stopping the repetition of dining eating places.

1.2. Tool-usage Module

TravelAgent integrates diverse actual-time records sources to make sure up to date recommendations and plans. Tool-utilization Module, which underpins Rationality and Personalization and serves because the middle issue for enhancing Comprehensiveness, is divided into classes: **actual-time API gear** and **arithmetic set of rules gear**. The first class, powered by means of SerpAPI2 and Google Maps API3, consists of specialized utilities such as Attraction Tool, designed to fetch tremendous actual- time statistics. The 2nd category comprises tools like Distance and Time Tools, which appoint mathematics algorithms to make sure specific course planning.

As proven in Figure 2, after initial constraints modeling, Tool-utilization Module extracts established parameters from Hard Constraints, serving because the foremost filter out at the device-name stage. Subsequently, it invokes specific tools according to the centered advice or planning factor. The filtered information is then transmitted to Recommendation Module for added refinement and in the end, conveyed to Planning Module.

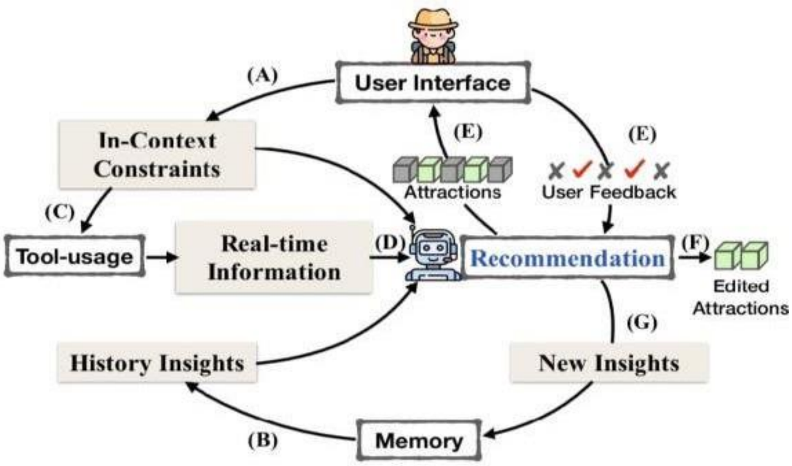

Figure 3: The process of Attraction Recommendation

1.3. Recommendation Module

Recommendation Module serves as the number one factor to gain Personalization of TravelAgent. It diverges from traditional paradigms of recommendation systems by means of at once leveraging LLM-based totally dealers to carry out online pointers. Inspired by way of some related paintings for on-line Personalization construction (Chidlovskii, 2015; Dalvi et al., 2022; Baek et al., 2024; Ma et al., 2024), we shape a data-driven recommendation framework incorporates Budget, City, Flight, Hotel, Restaurant, and Attraction Recommender, which always learns implicit personalization records thru more than one interactions.

Recommendation Module capitalizes on 4 awesome records resources for tips:

- **In-context Travel Constraints**, sourced from the person interface, refers to Hard Constraints supplied via customers, reflecting their specific wishes for the cutting-edge journey scenario.
- **Real-time Travel Information**, derived from **Tool- utilization Module**, encompasses up-to-date details on numerous travel options, every characterized through multi- dimensional attributes.
- **History Insights**, retrieved from **Memory Module**, contains discovered common-sense know-how, person personas, and beyond interactions facts. These insights dynamically evolve via new interactions and in the end re-storing in Memory Module for future use.
- **Generalized Knowledge**, sourced from **LLMs**, consists of a extensive variety of fundamental tour-related know-how.

Take appeal recommendation as an instance, the entire worflow is proven in Figure three **- (A) In-context Constraints Extraction:** At the begin of a journey scenario, Hard Constraints are initialized based at the consumer's enter from the user interface; **(B) History Insights Retrieval:** Concurrently, the device retrieves relevant historic insights from Memory

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.311 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 12, Issue 5, May 2025
DOI: 10.17148/IARJSET.2025.125334

Module, initializing implicit constraints, consisting of both soft and common sense constraints; **(C) Real-time Information Gathering:** At the enchantment advice degree, Attraction Tool makes use of Hard Constraints to fetch up to date attraction statistics;
(D) Data-driven Recommendation: Recommendation Module synthesizes all information and leverage LLMs to generate personalized enchantment recommendations and visit tips;
(E) User Behavioral Interactions: Recommendation list, followed with the aid of contextual facts, are presented on the established person interface for user comments (like/bypass); (F) **Items Re-rank** Based on consumer feedback, the Recommendation Module re-ranks the attractions, finalizing the choice for course making plans;
(G) Implicit Insight Learning: Post-interaction, the system draws new brief-time period in-points of interest applicable to users' personality and common-sense information, updating lengthy-time period insights in Memory Module. With this pipeline, TravelAgent can usually research implicit information and dynamically hold a tailor-made perception of each person.

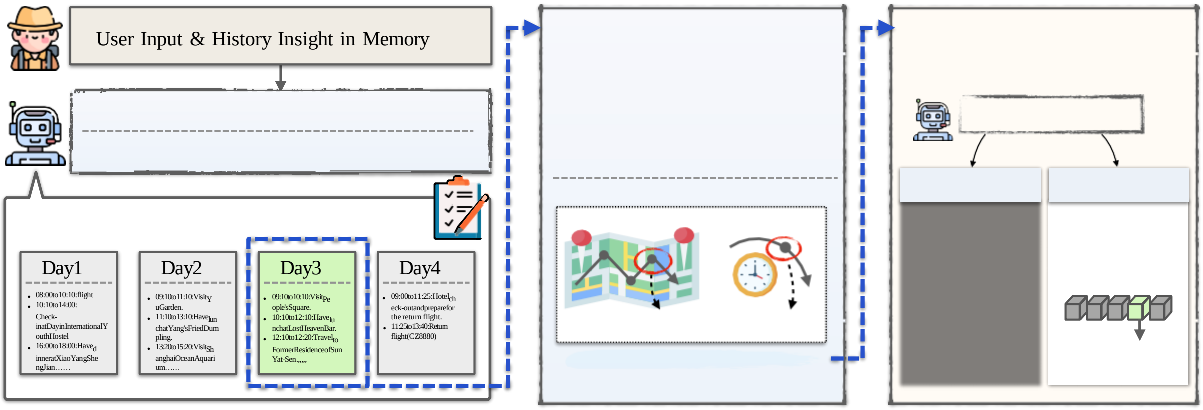

Figure 4 : The whole workflow of Route Planner illustrates the dynamic generation process of multi-day itineraries.

1.4. Planning Module

Planning Module accommodates **Budget Planner** and **Route Planner**. We innovatively integrate budgetary elements into the making plans manner to enhance the itinerary's Rationality and Comprehensiveness, empowering making plans with LLMs and a spatiotemporal-aware direction algorithm.

Budget Planner: After initial constraint modeling, Budget Recommender engages **LLMs** to generate a initial finances estimate. Once users verify their finances, Budget Planner re-evaluates customers' monetary situation and strategically distributes the budget throughout six classes: Accommodation, Attractions, Restaurants, Transportation, Other Expenses, and a Reserve Fund, primarily adheres to common sense understanding and dynamically adjusts consistent with different constraints to ensure that the travel plans are economically feasible and tailored to the context.

Route Planner: As proven in Figure four, guided with the aid of the finances plan, multi-day itineraries are dynamically generated the use of an appeal set seasoned- vided by using Recommendation Module. Each day's

Algorithm 1: Calculate Score for Attraction Ai
Require:
Current Position pcur , Current Time tcur
Available Attraction Ai = (pi, δti, τi)
▷ Attraction Position Ai
▷ Recommended Duration δti
▷ Recommended Visit Time Window τi
Left Time tlef t, **Travel Speed** v
Ensure: Si
1: di ← ‖pcur − pi‖ ▷ Calculate distance
2: tcost,i ← di/v + δtrec,i ▷ Calculate travel time 3: Tarr,i ← tcur + di/v ▷ Calculate arrival time 4: Tdep,i ← Tarr,i + δti ▷ Calculate leave time 5: Sret,i ← RETURNTIMESCORE(tlef t, tcost,i)
6: Sopt,i ← VISITTIMESCORE((Tarr,i, Tdep,i), τi) 7: Slef t,i ← RESERVETIMESCORE(tlef t, tcost,i) 8: Si ← Sret,i

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ꞔ Peer-reviewed & Refereed journal ꞔ Vol. 12, Issue 5, May 2025

DOI: 10.17148/IARJSET.2025.125334

+ Sopt,i + Slef t,i ▷ Final score
9: **return** Si

planning starts off evolved by means of setting initial conditions, including modern role pcur, modern time tcur, and an end time. In addition, we pre-described some expected occasion workouts consistent with Commonsense Constraints to make sure rational making plans, such as eating or check- ins. The day by day making plans manner iterates as long as time permits and points of interest are to be had, with every new release re-comparing the course point characterized with the aid of the modern coordinates (pcur, tcur). If the modern coordinates align with a specific event, corresponding actions are carried out, which includes selecting an foremost nearby eating place. Otherwise, the planner evaluates the available points of interest and selects one for a go to. After every path point, the coordinates (pcur,tcur) are updated to align subsequent activities with the strategic temporal and spatial logistics.

Attraction Scores Calculation: For higher assessment at (pcur, tcur), we devise Algorithm 1 to attain to be had points of interest. For every attraction Ai, firstly take a look at the space di from pcur to the enchantment pi to compute three time-(1) the travel time tcost,i, including both the tour length at speed v and the endorsed go to period δtrec,i; (2) the advent time Tarr,i (3) the go away time Tdep,i. The scoring mechanism includes 3 components:

- **Return Time Score** Sret,i, which assessments if the overall tour time and length at the Ai suit in the closing day time tleft.
- **Optimal Visit Time Score** Sopt,i, which evaluates how properly the visit aligns with the desired traveling window τi.
- **Reserve Time Score** Sleft,i, which optimizes the usage of the remaining time, prioritizing efficient day planning.

The general score Si = Sret,i +Sopt,i +Sleft,i guides the selection of the next enchantment, making sure the itinerary is balanced, enjoyable, and logistically sound. This approach dynamically adapts to each day instances, aligning tour plans with real-time constraints and choices.

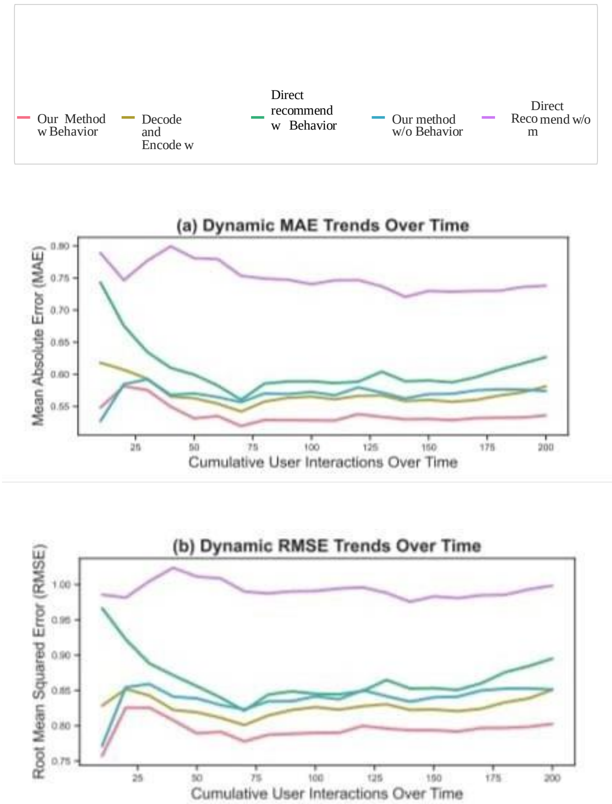

Figure 5: MAE and RMSE trends over simulated user interactions. When with user behavior history, our methods demonstrate superior cold-start performance and excellent error control. When without user behavior history, our methods mitigate performance degradation effectively compared to baselines.

DOI: 10.17148/IARJSET.2025.125334

1.5. Memory Module

Memory Module is designed to promote Personalization and dynamic records processing and includes two primary components: **(1) Short-time period Memory** stores new person interactions and captures the instant contextual tour statistics, constraints, and plans, adapting to the trendy person options. **(2) Long-term Memory** retains consumer statistics over prolonged intervals, accumulating a complete character based on historical interactions. It shops strong person possibilities, inclusive of a user's consistent hobby in green attractions or unique nutritional requirements. Historical insights permit the gadget to provide guidelines that align with long-time period tour behavior and alternatives.

III. EVALUATION

We conduct simulated and human user evaluations of TravelAgent to verify its performance against the three criteria: Rationality, Comprehensiveness, and Personalization.

Overall Evaluation

This section gives the outcomes from case studies that evaluated TravelAgent using 20 distinct travel scenarios and user input forms (proven in Appendix D). To establish a baseline, those case studies were additionally carried out with an ordinary agent device, the GPT-four+ agent (Achiam et al., 2023), to exhibit the overall performance of the proposed TravelAgent. In this section, human customers engage with TravelAgent and the GPT-4+ agent online the use of those 20 enter bureaucracy. Throughout the entire workflow and analysis of generated itineraries, human users evaluated 3 criteria on a scale from 1 to 10: Rationality, Comprehensiveness, and Personalization. Each criterion was assessed in keeping with a predefined set of assessment dimensions distinctive in Appendix E. The average ratings are supplied in Table 1, and comparative case study examples are designated in Appendices F and G

Personalized Evaluation

To examine Personalization Module of TravelAgent and the performance of our agent-based totally on-line advice framework, we test with a hundred simulated customers based on GPT-3.5-rapid-0125 (OpenAI, 2022) and 10 travel situations and teach Recommendation Module to signify 20 points of interest consistent with city to each user. Simulated customers charge items primarily based on alternatives, and TravelAgent predicts those ratings whilst usually mastering over interactions. We in comparison our technique with two famous LLM-based recommendation frameworks: **(1) Direct Recommend** and **(2) Decode and Encode**, both lacking reminiscence and chronic getting to know. Evaluation metrics included Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). Figure 5 indicates the trend results over consumer interactions and our methods' high prediction accuracy, verifying TravelAgent's Personalization.

Limitations

While TravelAgent we proposed goals to address rational and personalized journey itinerary era for actual-world tour situations, its miles crucial to acknowledge sure obstacles in our have a look at. Firstly, the machine closely is predicated on the first-class of dynamic real-time records resources; inaccuracies, which include previous flight information, can bring about suboptimal recommendations and impractical final plans. Future research should recognition on improving the exceptional and reliability of these statistics assets to improve adaptability throughout exclusive travel scenarios. Secondly, a deeper exploration of personalized sports in journey itineraries based on extra targeted consumer personalization records is needed for in addition improvement.

IV. CONCLUSION

In this work, we introduced TravelAgent, a singular journey service gadget leveraging LLMs and real- time information for automatic itinerary technology in actual- international scenarios with multi-dimensional constraints. We set up three criteria for travel planning and version tour constraints. To deal with the challenges, we evolved tailored spatiotemporal-conscious algorithms and a novel agent-based customized on-line recommendation framework better with the aid of real-time records processing mechanism and dynamic memory. We examine TravelAgent via simulated user rating prediction and actual-international human case studies. The system promises rational, comprehensive, and personalized journey offerings. We desire our work will encourage in addition exploration of travel services and LLM-based agent programs.

REFERENCES

[1]. Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman,Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, et al. 2023. Gpt-4 technical report. arXiv preprint arXiv:2303.08774.
[2]. Jinheon Baek, Nirupama Chandrasekaran, Silviu Cucerzan, Allen Herring, and Sujay Kumar Jauhar. 2024. Knowledge-augmented large language models for personalized contextual query suggestion. In Proceedings of the

DOI: 10.17148/IARJSET.2025.125334

ACM on Web Conference 2024, pages 3355–3366.

[3]. Maciej Besta, Nils Blach, Ales Kubicek, Robert Gerstenberger, Michal Podstawski, Lukas Gianinazzi, Joanna Gajda, Tomasz Lehmann, Hubert Niewiadomski, Piotr Nyczyk, et al. 2024. Graph of thoughts: Solving elaborate problems with large language models. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 38, pages 17682–17690.

[4]. Booking Holdings Inc. 2024. Booking.com | official site | the best hotels, flights, car rentals & accommodations. https://www.booking.com/. Accessed on August 5, 2024.

[5]. Wenhu Chen, Xueguang Ma, Xinyi Wang, and William W Cohen. 2022. Program of thoughts prompting: Disentangling computation from reasoning for numerical reasoning tasks. arXiv preprint arXiv:2211.12588.

[6]. Boris Chidlovskii. 2015. Improved trip planning by learning from travelers' choices. In Proceedings of the 2nd International Conference on Mining Urban Data-Volume 1392, pages 17–26.

[7]. Sunhao Dai, Ninglu Shao, Haiyuan Zhao, Weijie Yu, Zihua Si, Chen Xu, Zhongxiang Sun, Xiao Zhang, and Jun Xu. 2023. Uncovering chatgpt's capabilities in recommender systems. In Proceedings of the 17th ACM Conference on Recommender Systems.

[8]. Bhavana Dalvi, Oyvind Tafjord, and Peter Clark. 2022. Towards teachable reasoning systems: Using a dynamic memory of user feedback for continual system improvement. In Proceedings of the 2022 conference on empirical methods in natural language processing, pages 9465–9480.

[9]. Expedia Group. 2024. Expedia travel: Vacation homes, hotels, car rentals, flights & more. https://www. expedia.com/. Accessed on August 5, 2024.

[10]. FrommerMedia LLC. 2024. Frommer's travel guides: Trip ideas, inspiration & deals. https://www. frommers.com/. Accessed on August 5, 2024.

[11]. Masum Hasan, Cengiz Ozel, Sammy Potter, and Ehsan Hoque. 2023. Sapien: affective virtual agents powered by large language models. In 2023 11th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW), pages 1–3. IEEE.

[12]. Yupeng Hou, Junjie Zhang, Zihan Lin, Hongyu Lu, Ruobing Xie, Julian McAuley, and Wayne Xin Zhao. 2024. Large language models are zero-shot rankers for recommender systems. In European Conference on Information Retrieval, pages 364–381. Springer.

[13]. Chenxu Hu, Jie Fu, Chenzhuang Du, Simian Luo, Junbo Zhao, and Hang Zhao. 2023. Chatdb: Augmenting llms with databases as their symbolic memory. arXiv preprint arXiv:2306.03901.

[14]. Ziheng Huang, Sebastian Gutierrez, Hemanth Kamana, and Stephen MacNeil. 2023. Memory sandbox: Transparent and interactive memory management for conversational agents. In Adjunct Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology, pages 1–3.

[15]. Albert Q Jiang, Alexandre Sablayrolles, Arthur Mensch, Chris Bamford, Devendra Singh Chaplot, Diego de las Casas, Florian Bressand, Gianna Lengyel, Guillaume Lample, Lucile Saulnier, et al. 2023. Mistral 7b. arXiv preprint arXiv:2310.06825.

[16]. Albert Q Jiang, Alexandre Sablayrolles, Antoine Roux, Arthur Mensch, Blanche Savary, Chris Bamford, Devendra Singh Chaplot, Diego de las Casas, Emma Bou Hanna, Florian Bressand, et al. 2024. Mixtral of experts. arXiv preprint arXiv:2401.04088.

[17]. Wang-Cheng Kang, Jianmo Ni, Nikhil Mehta, Maheswaran Sathiamoorthy, Lichan Hong, Ed Chi, and Derek Zhiyuan Cheng. 2023. Do llms understand user preferences? evaluating llms on user rating prediction. arXiv preprint arXiv:2305.06474.

[18]. Junling Liu, Chao Liu, Peilin Zhou, Renjie Lv, Kang Zhou, and Yan Zhang. 2023a. Is chatgpt a good recommender? a preliminary study. arXiv preprint arXiv:2304.10149.

[19]. Lei Liu, Xiaoyan Yang, Yue Shen, Binbin Hu, Zhiqiang Zhang, Jinjie Gu, and Guannan Zhang. 2023b. Think-in-memory: Recalling and post-thinking enable llms with long-term memory. arXiv preprint arXiv:2311.08719.

[20]. Pan Lu, Baolin Peng, Hao Cheng, Michel Galley, KaiWei Chang, Ying Nian Wu, Song-Chun Zhu, and Jianfeng Gao. 2024. Chameleon: Plug-and-play compositional reasoning with large language models. Advances in Neural Information Processing Systems, 36.

[21]. Xiao Ma, Swaroop Mishra, Ariel Liu, Sophie Ying Su, Jilin Chen, Chinmay Kulkarni, Heng-Tze Cheng, Quoc Le, and Ed Chi. 2024. Beyond chatbots: Explore llm for structured thoughts and personalized model responses. In Extended Abstracts of the CHI Conference on Human Factors in Computing Systems, pages 1–12.

[22]. OpenAI. 2022. Chatgpt. https://openai.com/blog/ chatgpt . Accessed: 2024-08-06.

[23]. Yujia Qin, Shihao Liang, Yining Ye, Kunlun Zhu, Lan Yan, Yaxi Lu, Yankai Lin, Xin Cong, Xiangru Tang, Bill Qian, et al. 2023. Toolllm: Facilitating large language models to master 16000+ real-world apis. arXiv preprint arXiv:2307.16789.

[24]. Roadtrippers, LLC. 2024. Road trip planner – build your itinerary and find the best stops. https:// roadtrippers.com/. Accessed on August 5, 2024.

[25]. Timo Schick, Jane Dwivedi-Yu, Roberto Dessì, Roberta Raileanu, Maria Lomeli, Eric Hambro, Luke

DOI: 10.17148/IARJSET.2025.125334

Zettlemoyer, Nicola Cancedda, and Thomas Scialom. 2024. Toolformer: Language models can teach themselves to use tools. Advances in Neural Information Processing Systems, 36.

[26]. Yongliang Shen, Kaitao Song, Xu Tan, Dongsheng Li, Weiming Lu, and Yueting Zhuang. 2024. Hugging gpt: Solving ai tasks with chatgpt and its friends in hugging face. Advances in Neural Information Processing Systems, 36.

[27]. Noah Shinn, Federico Cassano, Ashwin Gopinath, Karthik Narasimhan, and Shunyu Yao. 2024. Reflexion: Language agents with verbal reinforcement learning. Advances in Neural Information Processing Systems, 36.

[28]. Ishika Singh, Valts Blukis, Arsalan Mousavian, Ankit Goyal, Danfei Xu, Jonathan Tremblay, Dieter Fox, Jesse Thomason, and Animesh Garg. 2023. Prog-prompt: Generating situated robot task plans using large language models. In 2023 IEEE International Conference on Robotics and Automation (ICRA), pages 11523–11530.IEEE.

[29]. Melanie Swan, Takashi Kido, Eric Roland, and Renato P dos Santos. 2023. Math agents: Computational infrastructure, mathematical embedding, and genomics. arXiv preprint arXiv:2307.02502.

[30]. G Team, Anil R, S Borgeaud, et al. 2023. Gemini: a family of highly capable multimodal models. arXiv preprint arXiv:2312.11805.

[31]. Boshi Wang, Hao Fang, Jason Eisner, Benjamin Van Durme, and Yu Su. 2024a. Llms in the imaginarium: tool learning through simulated trial and error. arXiv preprint arXiv:2403.04746.

[32]. Lei Wang and Ee-Peng Lim. 2023. Zero-shot next-item recommendation using large pretrained language models. arXiv preprint arXiv:2304.03153.

[33]. Lei Wang, Chen Ma, Xueyang Feng, Zeyu Zhang, Hao Yang, Jingsen Zhang, Zhiyuan Chen, Jiakai Tang, Xu Chen, Yankai Lin, et al. 2024b. A survey on large language model based autonomous agents. Frontiers of Computer Science, 18(6):186345

[34]. Lei Wang, Wanyu Xu, Yihuai Lan, Zhiqiang Hu, Yunshi Lan, Roy Ka-Wei Lee, and Ee-Peng Lim. 2023. Plan-and solve prompting: Improving zero-shot chain-of-thought reasoning by large language models. arXiv preprint arXiv:2305.04091.

[35]. Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed Chi, Quoc V Le, Denny Zhou, et al. 2022. Chain-of-thought prompting elicits reasoning in large language models. Advances in neural information processing systems, 35:24824–24837.

[36]. Wikitravellers. 2024. Wikitravel - the free travel guide. https://wikitravel.org/.

[37]. Jian Xie, Kai Zhang, Jiangjie Chen, Tinghui Zhu, Renze Lou, Yuandong Tian, Yanghua Xiao, and Yu Su. 2024. Travelplanner: A benchmark for real-world planning with language agents. arXiv preprint arXiv:2402.01622.

[38]. Fan Yang, Zheng Chen, Ziyan Jiang, Eunah Cho, Xiaojiang Huang, and Yanbin Lu. 2023. Palr: Personalization aware llms for recommendation. arXiv preprint arXiv:2305.07622.

[39]. Shunyu Yao, Dian Yu, Jeffrey Zhao, Izhak Shafran, Tom Griffiths, Yuan Cao, and Karthik Narasimhan. 2024. Tree of thoughts: Deliberate problem solving with large language models. Advances in Neural Information Processing Systems, 36.

[40]. Shunyu Yao, Jeffrey Zhao, Dian Yu, Nan Du, Izhak Shafran, Karthik Narasimhan, and Yuan Cao. 2022. React: Synergizing reasoning and acting in language models. arXiv preprint arXiv:2210.03629.

[41]. Siyu Yuan, Kaitao Song, Jiangjie Chen, Xu Tan, Yongliang Shen, Ren Kan, Dongsheng Li, and Deqing Yang. 2024. Easytool: Enhancing llm-based agents with concise tool instruction. arXiv preprint arXiv:2401.06201.

[42]. Hongbo Zhang, Junying Chen, Feng Jiang, Fei Yu, Zhi-hong Chen, Jianquan Li, Guiming Chen, Xiangbo Wu, Zhiyi Zhang, Qingying Xiao, et al. 2023. Huatuo gpt, towards taming language model to be a doctor. arXiv preprint arXiv:2305.15075.

[43]. Longtao Zheng, Rundong Wang, Xinrun Wang, and Bo An. 2023. Synapse: Trajectory-as-exemplar prompting with memory for computer control. In The Twelfth International Conference on Learning Representations.

[44]. Wanjun Zhong, Lianghong Guo, Qiqi Gao, He Ye, and Yanlin Wang. 2024. Memorybank: Enhancing large language models with long-term memory. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 38, pages 19724–19731.

[45]. Joyce Zhou, Yijia Dai, and Thorsten Joachims. 2024. Language-based user profiles for recommendation. arXiv preprint arXiv:2402.15623.