



Charles W. Davidson College of Engineering (CoE)
Fall 2021 Davidson Student Scholar Proposal

Student Name (Last, First): **Vijayakumar, Sudha**
SJSU Student ID#: 014749488
Email Address: sudha.vijayakumar@sjsu.edu
Grade Point Average (at start of Fall 2021): 3.95

CoE Department: Charles W. Davidson College of Engineering
Undergraduate or Graduate & Year: 2nd year Graduate, 2021-22
Expected Graduation Semester/Year: Spring 2022

Proposed Research Type ('X' one): ____ Project **X** Thesis (Graduate Students only)

Estimated Number of Hours You Expect to Devote to Research per Week: 20 hours/ week

Thesis/Project Title (< 10 words)

SmartRec - Smart Conversational Recommendation System using Semantic Knowledge Graphs

Synopsis of Proposed Research (≤ 250 words)

Conversational Recommendation Systems (CRS) are gaining popularity as a next-generation enabler of e-commerce systems. A CRS is a more natural way for users to interact with e-commerce platforms[3] and mimics the way users interact with customer support personnel or a knowledgeable friend in everyday life. The need for CRS became even more pronounced during the unprecedented COVID-19 pandemic[11] when customer support systems became unavailable and left customers unable to ask questions or resolve issues leading to widespread deployment of basic chatbots. We envision building a smart CRS that can provide product recommendations & customer support in real-time through an interactive conversation with the user. Crucial challenges that we need to overcome in our proposed system that most state-of-the-art approaches suffer is to address the lack of contextual understanding of individual user preferences (e.g., what is the user's current goal when booking a vacation), bridge the

semantic gap between user preferences in natural language expressions and product attribute definitions. To overcome these issues, we will build a *context-aware semantic knowledge graph* that captures the current conversation, past conversations, and product attributes to produce relevant and semantically meaningful responses to user queries. We will then add a product recommendation and dialogue components that rely on this knowledge graph. As part of our study, we will evaluate machine learning, system performance, and user experience and compare against state-of-the-art CRS's (e.g., [4,10]).

CoE Faculty Advisor/Mentor: Name (Last, First): Jetcheva, Jorjeta

I have reviewed the student's proposal and agree that the proposed tasks include a research component of originality and scope adequate for consideration for a monetary award upon completion. I agree to supervise the proposed research until completion.

Faculty Signature:

Date: 09/27/2021

A handwritten signature in cursive script, appearing to read "Jorjeta Jetcheva".

Body of the Proposal

A. Introduction

A Forbes article states that 72% of E-commerce users prefer to seek answers online. By 2022, according to Gartner reports, 70% of middle-class workers will utilize conversational stages each day [1]. 58% of consumers say that they care even more than before about customer experience in the post COVID world [2]. Covid-19 pandemic has undoubtedly accelerated the pace at which E-commerce platforms will adapt conversational agents for highly available customer support and, Gartner forecasts second that. This projected demand calls for the need to implement more generic and efficient conversational agents & recommendation engines that can provide customers with the required information at every stage of the conversation during the purchasing and issue-resolving cycle. In this proposal, we present SmartRec, a Smart Conversational AI-based recommendation system. SmartRec can have a multi-turn conversation with the user to understand the context and semantics behind their product requirement. SmartRec can generate appropriate recommendations/ natural language response text suggestions based on the user queries. Most of the state-of-the-art CRS are domain-specific, lack meaningful multi-turn conversation, lack contextual understanding of the user preferences, and possess a semantic gap between user intent in natural language and product attributes. To resolve these issues, we propose a solution that offers (1) an auto-suggest type-ahead mechanism [5] to help users convey their intent driving for more meaningful conversations (2) semantic knowledge graphs that captures the context of the user intent in natural language and map it to the corresponding products and (3) an integrated recommendation & dialog engine to generate product and solution recommendations using the current and historical user interactions. The goal is to fuse current data from the user conversation, historical user-item interactions, historical issue handling data, and social-conversational data of a particular brand to better understand the user, match and, return the products/ solutions that interest them.

B. Specific Proposed Research and Why Important/Exciting

Through our work, we propose to implement an expert CRS using publicly available large-scale Airbnb data. The large-scale Airbnb data presents a substantial modeling challenge for our proposed study and stands as a representative of a broad range of use cases involving consumer choices and decision-making. In turn, this shall reflect the notion that the models resulting from our work will have applicability for customer support in the context of similar e-commerce or vacation planning applications. We plan to leverage Random Access Navigation(RAN) [8] type-ahead approach-based UI to drive a freestyle yet guided conversation for the user. The type-ahead approach allows users to converse with guidance and make corrections anytime as humans act in real life. Probabilistically, RAN only looks for missing parameters and then asks only relevant questions eliminating the need to hard-code all the possible $n*(n-1)$ cases for the given n parameters. A good product recommendation or customer support system relies upon an accurate interpretation of the user preferences. Therefore, we propose to leverage the semantic

fusion approach to contextually map the user preferences to the available products or suggestions while remembering the collective requirement from the chat history. A word-oriented knowledge graph [13] (i.e., ConceptNet3) using Airbnb's social conversational data, issue handling data, and a custom product-attribute knowledge graph using Airbnb's item listings, historical user interactions, and user reviews will add data intelligence to the system. We plan to include a neural collaborative filtering framework [6] to implement an integrated recommendation cum dialog component [7] that will act as a retrieval & generative response model to list the relevant recommendations and response text. We intend to implement a generic solution architecture for an AI-enhanced product recommendation & customer support systems in E-Commerce.

C. Methodology and Why Innovative/Creative

State-of-the-art knowledge graph-based CRS's [4] trained on large-scale multi-domain data but do not generalize well for a specific domain [13], lack features that can assist interactive conversations, and do not include historical user-item interactions. We envision building a novel CRS utilizing semantic knowledge graph and machine learning components, incorporating a hybrid response model which combines the ability to list product recommendations with suggestions for the reported customer issue. As part of our research, we will curate a large-scale, domain-relevant, integrated dataset from different sources such as Airbnb tagged data from Twitter & Quora, public data released by Airbnb, including listings, user reviews, FAQs. Our CRS will incorporate preference understanding by current context, historical user interactions, and user engagement.

We intend our proposed system architecture to be generic and thus easy to train with data for different domains and adapted to a broad range of AI-enhanced customer service applications. Future extensions of this solution will incorporate image and video interactions that accompany the conversation, voice-based, and multi-lingual CRS [9], expanding the potential footprint of the solution to an even a broad range of applications and users.

D. Milestones and Timeline

Task	Target Completion
Data scraping, storage & preprocessing	October 10, 2021
Construct knowledge graphs - Investigate & Implement best approaches	November 1, 2021
UI, Recommendation & Dialog Components	December 1, 2022
Training & Testing using different pre-defined Personas	February 1, 2022
Deployment & Human Evaluation	April 15, 2021
Prepare thesis	May 15, 2021
Submit a manuscript for conference proceedings/journal	May 20, 2021

E. Anticipated Outcome

We intend the proposed system to be deployed to many E-commerce platforms that include big players like Airbnb, Amazon to enhance customer experience and provide an AI-enabled customer recommendation and support system. This study will focus on an Airbnb-inspired use case for which a large amount of data has been made available to the public by the company.

Airbnb has implemented a task-oriented customer support bot [12] to provide customer support during the COVID pandemic and has added flexible filters to tackle the rigidity of the current product search system. SMARTRec has the potential to significantly expand these existing capabilities with a much more sophisticated AI-based approach. On thesis completion, we intend to present this work to Airbnb and submit the paper to the ACL 2022 conference.

F. Budget

This project shall require the use of a College of Engineering HPC(High-Performance Computing) environment. For additional experiments, we are planning to use 40 free GPU hours per week offered by kaggle. We anticipate the initial storage requirement to be around 200GB and planning to use the HPC & Google Drive(Student account with unlimited data) to store all relevant data and results. Therefore, with all the free-of-cost infrastructure in place, we estimate the budget to be ‘zero cost’ at this point.

G. References

- [1] Abhimanyu Singh, ‘*Harnessing Conversational Voice AI In The E-Commerce Industry*’, *Forbes*, August 24, 2021. [Accessed: September 17, 2021]
- [2] Brenda Hodge, ‘*Kicking the digital front door open*’, *What’s next*, August 3, 2021. [Accessed: September 17, 2021]
- [3] Zhang, Y., Chen, X., Ai, Q., Yang, L., & Croft, W. B. (2018, October). Towards conversational search and recommendation: System ask, user respond. In *Proceedings of the 27th acm international conference on information and knowledge management* (pp. 177-186).
- [4] Zhou, K., Zhao, W. X., Bian, S., Zhou, Y., Wen, J. R., & Yu, J. (2020, August). Improving conversational recommender systems via knowledge graph based semantic fusion. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 1006-1014).
- [5] Li, G., Wang, J., Li, C., & Feng, J. (2012, August). Supporting efficient top-k queries in type-ahead search. In *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval* (pp. 355-364).
- [6] Sang, L., Xu, M., Qian, S., & Wu, X. (2021). Knowledge graph enhanced neural collaborative recommendation. *Expert Systems with Applications*, 164, 113992.
- [7] Hedayatnia, B., Gopalakrishnan, K., Kim, S., Liu, Y., Eric, M., & Hakkani-Tur, D. (2020). Policy-Driven neural response generation for Knowledge-Grounded dialogue systems. *arXiv preprint arXiv:2005.12529*.
- [8] Shane Mac, ‘*How to use Random Access Navigation (RAN) to create smarter bots*’, *medium*, June 29, 2017. [Accessed: September 17, 2021]
- [9] Havasi, C., Speer, R., & Alonso, J. (2007, September). ConceptNet 3: a flexible, multilingual semantic network for common sense knowledge. In *Recent advances in natural language processing* (pp. 27-29). Philadelphia, PA: John Benjamins.
- [10] Wong, C. M., Feng, F., Zhang, W., Vong, C. M., Chen, H., Zhang, Y., ... & Chen, H. (2021, April). Improving Conversational Recommender System by Pretraining Billion-scale Knowledge Graph. In *2021 IEEE 37th International Conference on Data Engineering (ICDE)* (pp. 2607-2612). IEEE.
- [11] Chris Zhu, ‘*Using Chatbots to Provide Faster COVID-19 Community Support*’, *medium*, July 15, 2020. [Accessed: September 17, 2021]

- [12] Marta Lamothe, '*Airbnb's 100+ new app features: Everything you need to know*', *your.rentals*, June 7, 2021. [Accessed: September 17, 2021]
- [13] Havasi, C., Speer, R., & Alonso, J. (2007, September). ConceptNet 3: a flexible, multilingual semantic network for common sense knowledge. In *Recent advances in natural language processing* (pp. 27-29). Philadelphia, PA: John Benjamins.