7 Conclusion

In this final chapter of the thesis we provide a brief summary of the most important aspect and insights. We also go over benefits of our results but also limitations where our work does not hold. Lastly we mention lessons learned and additionally suggestions for future work.

7.1 Summary

This thesis explored the development and evaluation of an innovative chatbot system designed to enhance user interaction and system explainability in smart home environments. The research focused on integrating LLMs without fine-tuning to support diverse intents within the Bosch Smart Home ecosystem, addressing the growing complexity of smart home technologies.

A novel approach combining natural language processing with JSON function calling was implemented, enabling the chatbot to handle both device control and data analysis tasks. The system architecture, utilizing a client-server model with Android Studio and Ollama, demonstrated practical deployment potential in real-world scenarios. However, for a production environment, tools like LangChain offer additional features like scalability and deployment tools that are crucial for handling larger user bases and simplifying the deployment process (it also supports the use of Ollama). The evaluation process introduced a comprehensive metric that combined semantic similarity and JSON accuracy, providing a nuanced assessment of the chatbot's performance. This revealed interesting trade-offs between JSON accuracy and semantic understanding across different model iterations. The `shllama3instruct` model (as we named our customized model) emerged as the top performer, balancing high semantic similarity with good precision in JSON function calling.

User studies validated the chatbot's positive impact on system explainability and usability. However, challenges remained in user preference for the chatbot over traditional interfaces and in optimizing device control task efficiency. Qualitative feedback highlighted the importance of balancing simplicity with advanced functionality and addressing privacy concerns in smart home applications.

The research uncovered potential for developing a generalized framework for smart home chatbots, which could significantly impact future developments in smart home technology. Areas for improvement were identified, including enhanced contextual understanding, integration of multimodal inputs or even outputs, better customization of the language model to always answer in the language of the last user message and the addressing of complex use cases, especially on the topic energy consumption.

This thesis contributes to the field by demonstrating the feasibility and benefits of integrating recent LLMs into smart home systems. It offers valuable insights into the challenges and opportunities in enhancing user interaction with complex smart home ecosystems, paving the way for more intuitive and explainable smart home technologies.

7.2 Benefits

The thesis results benefit various stakeholders in the smart home and software development domains. Software architects and developers gain insights into open-source tools and building blocks for implementing chatbots in smart home systems. Individuals interested in customizing models without fine-tuning can leverage the approach demonstrated. LLM researchers benefit from performance data on models within the tested size range, providing valuable benchmarks for future studies. Companies like Bosch Smart Home, focused on maintaining high customer data security and privacy, can utilize the approach that keeps customer data within the company by running customized LLMs on internal servers. This addresses the crucial balance between functionality and data protection. Additionally, the research offers insights into which data types are most relevant for smart home chatbots, helping companies prioritize data collection and usage. Overall, the thesis contributes to advancing smart home technology while addressing key industry concerns, making it valuable for both technical professionals and business stakeholders in the smart home sector.

7.3 Limitations

The study's limitations primarily stem from its focused scope and methodological choices. Firstly, the exclusive use of open-source language models means that the performance of commercial or proprietary models remains unexplored, potentially offering different results. The function calling approach, centered on mapping JSONs to functionality, doesn't account for newer methods available in select models like Mistral. This limitation raises questions about the approach's effectiveness in other domains beyond smart homes. Additionally, the specialization of data and focus on smart home applications may limit the generalizability of findings to other fields. Lastly, the study doesn't address the scalability of the approach for larger, more complex smart home systems with a wider range of devices and functionalities.

7.4 Lessons Learned

The thesis process yielded valuable insights for future research endeavors. A key lesson was the time-intensive nature of multi-faceted evaluation approaches. While providing comprehensive insights, such methods demand significant resources. Future work might benefit from focusing on specific aspects or dividing the research into separate studies - one concentrating on system architecture and model performance, another on user-centric aspects. The manual creation of evaluation or training datasets proved exceptionally time-consuming, highlighting the need for more efficient data generation methods. The existence of tools for monitoring and regularly evaluating LLM models (for Ollama), such as ollama-grid-search[1] and promptfoo[2], was noted as potentially beneficial for future studies to accelerate and enhance the process. Lastly, the importance of

---

[1] https://github.com/dezoito/ollama-grid-search
[2] https://www.promptfoo.dev/

carefully selecting evaluation metrics for LLM use cases became evident. Allocating sufficient time for metric selection and implementation is crucial for ensuring the relevance and accuracy of research outcomes in this rapidly evolving field.

7.5 Future Work

There are various directions in which future research could go. Exploring larger parameter models or advanced fine-tuning/training approaches could enhance both JSON accuracy and semantic similarity. Developing more robust generalization capabilities would enable handling a wider range of intents and user requests. Implementing a more readable format for device data presentation to the model could potentially improve performance, especially for models other than llama3.

For Bosch Smart Home specifically, incorporating system logs and historical data functionality would be beneficial, necessitating research into effective data presentation and filtering methods for chatbots. Developing and evaluating features for energy optimization and system behavior insights could add significant value to the Bosch Smart Home system.

Investigating the impact of visual elements and voice interaction capabilities on user acceptance is another area that could be researched. Further, conducting larger-scale user studies with diverse participants and smart home setups would provide more comprehensive insights.

Finally, future research could also explore and categorize the various approaches to implementing AI across different applications, including different training methodologies, customization techniques (such as optimizing system messages, parameters, and examples for zero-/one-/few-shot learning), and the feasibility of building custom models from scratch, to determine the most effective and efficient methods for specific use cases and domains.