CAMARINES SUR POLYTECHNIC COLLEGES
COLLEGE of COMPUTER STUDIES

CCCS 106 - Application Development and Emerging Technologies

Project CCeSca: Conversational AI Assistant for Computer Science Learning and Project Development

Christine Kyla Belano
College of Computer Studies
Camarines Sur Polytechnic Colleges
chbelano@my.cspc.edu.ph

Donn Robert De Lima
College of Computer Studies
Camarines Sur Polytechnic Colleges
dodelima@my.cspc.edu.ph

Lily Rose Julianes
College of Computer Studies
Camarines Sur Polytechnic Colleges
lijulianes@my.cspc.edu.ph

1 Executive Summary

Project CCeSca (pronounced as 'Cheska') is a conversational AI assistant aimed at bridging the gap between theoretical knowledge and practical application in computer science education. Specifically designed for students in the College of Computer Studies, CCeSca uses artificial intelligence and machine learning to provide an interactive learning platform. It helps students navigate complex concepts, develop critical thinking skills, and foster collaborative learning. By leveraging open-source tools like NLTK and Flask, CCeSca offers cost-effective and flexible NLP and voice interaction features. Unlike enterprise-level platforms, CCeSca focuses on educational needs and affordability, making it accessible for students who are currently equipping themselves the knowledge and skills to thrive in the market.

To minimize costs, CCeSca will be hosted on a personal server, reducing hosting fees, while initial development expenses are projected to range between PHP 14,000 and PHP 18,000 for the first year. A freemium model will be adopted to generate revenue, with premium subscriptions priced at PHP 150 per month. With 5% of users projected to opt for the premium tier, CCeSca is expected to generate PHP 90,000 annually, covering operational costs and yielding a modest profit. This makes CCeSca a sustainable and scalable solution for enhancing student learning and addressing the growing demand for tech professionals equipped with practical skills.

2 Market Analysis

The chatbot market is dominated by several large competitors, including Dialogflow, Amazon Lex, and IBM Watson. These platforms offer Natural Language Processing (NLP) and voice interaction

capabilities but are often designed for broader, enterprise level applications making them less suitable for individual developers. Additionally, the cost of using this platforms along with their reliance on specific ecosystems can be prohibited for smaller-scale projects.

CCeSca stands out compared to the mentioned competitors as it offers specialized chatbot solution focused on computer related education. Laveraging open-source tools like NLTK, BERT, and Flask provides cost-effective flexible NLP and voice interaction capabilities tailored to its domain. Additionally, the development model used which is the Agile Development model allows the chatbot to improve its functionality and features over time as it can evolve iteratively based on the user feedback. The target expertise, advanced NLP, and the customization makes CCeSca a strong competitor in the educational chatbot market, Due to its use of free open source frameworks, a focused feature set, and customization flexibility, making it ideal for users seeking for a tailored, budget-friendly solution.

3 Technical Description

3.1 Franeworks and Libraries:

- Flask: A micro web framework that will serve as the foundation for the application's backend.
- NLTK: The Natural Language Toolkit will be utilized to process human language data, essential in the chatbot's NLP capabilities.

3.2 Core functionalities:

- Natural Language Processing: Enables the chatbot to understand and process user queries.
- Rule-Based Systems: Implements a dialogue management algorithm to guide conversations.
- BERT: Utilizes Bidirectional Encoder Representations from Transformers for effective response generation.

3.3 Voice Interaction:

- Web Speech API: Facilitates speech-to-text functionality for user input.
- SpeechSynthesis: Converts text responses generated by the application into spoken words, enhancing interactivity

3.4 User Interface:

- HTML, CSS, Bootstrap: These technologies will create a responsive and user-friendly interface for the application.

3.5 Backend Infrastructure

- MySQL: A relational database to store user data and interactions securely.

3.6 Machine Learning Integration:

- Recurrent Neural Network (RNN) | Long Short-Term Memory (LSTM): These advanced neural network architectures will improve the chatbot's ability to learn from sequential data and enhance the response accuracy.

4 Development Plan

The development of CCeSca chatbot will be structured according to the Agile Process Model, emphasizing iterative development, flexibility, and continuous improvement. This approach ensures that we can adopt to changes and continuously enhance the chatbot's capability. This project will be divided into following phases.

4.1 Planning:

This initial phase will define the project's scope and requirements, This involves identifying and documenting the key features and objectives of the chatbot to ensure the alignment with educational goals.

4.2 Design:

In this phase, the focus shifts to developing the architectural design of the project which includes designing a user interface that prioritize accessibility and usability, planning the backend infrastructure that supports chatbot operations as well as integrating machine learning models and related technologies. For initial review, a prototypes will be created to gather reviews and feedback to ensure that the design meets the need of the users. (not sure)

4.3 Development:

During development phase, natural language processing (NLP) capabilities and rule-based dialogue system will be developed and the integration of machine learning models particularly Long Short-Term Memory (LSTM) networks and Bidirectional Encoder Representations from Transformers(BERT) is also included as this phase focuses on implementing the chatbot's core functionality. In addition, user interaction flows will also be created to facilitate voice-based inputs and responses.

4.4 Testing:

In testing phase, conducting of unit tests followed by user testing will be performed in order to gather feedback on the project's performance and usability, to ensure it meets the user's expectations.

4.5 Iteration and Deployment:

Based on the received feedback, necessary improvements and additional features will be added before the chatbot is officially deployed to make sure the system meets the user's preference.

4.6 Post-Launch Monitoring:

The performance of CCesCa will be continuously monitored even after the initial launch, in connection with this, further enhancements will be implemented as a response to user feedback and performance data.

5 Financial Projection

5.1 Cost Estimation

To minimize expenses, Project CCeSca will be hosted on a personal server, eliminating hosting fees. Electricity and hardware maintenance are estimated at PHP 1,000 annually. Development tools might cost PHP 1,000 - 3,000, though open-source alternatives could reduce this to zero. Assuming 100 hours of team labor at PHP 100/hour, development costs would reach PHP 10,000. Operational costs, including updates and maintenance, add PHP 3,000 - 5,000 yearly. Overall, the first-year costs are expected to total **PHP 14,000 - 18,000**.

5.2 Revenue Projections

A freemium model will be adopted, with free basic features and paid premium subscriptions priced at PHP 150 per month. If 5% of 1,000 users opt for premium, this will generate **PHP 90,000 annually** (50 users × PHP 1,800/year). This revenue comfortably covers the initial costs, yielding a modest profit margin.

6 Marketing and Promotion Strategy

Marketing strategy for CCeSca will target educational institutions, particularly colleges and universities that offers computer related programs. The approaches to be employed are:

- Social Media Marketing: Utilizing various social media platforms such as Facebook, X, and YouTube to promote the product's functionality by engaging educational content and tutorials
- Partnership: We can establish collaborations with different educational platforms, colleges, and technology blogs to enhance visibility and reach among students and educators.
- Search Engine Optimization (SEO): Optimizing the CCeSca website and blog content can also attract organic traffic from students who are taking computer related programs as well as the educators that are searching for innovative educational tools.

7 Legal and Regulatory Promotion Strategy

CCeSca's development and operation will strictly adhere to the Philippine Data Privacy Act of 2012 (RA 10173), ensuring compliance with regulations governing data collection, user consent, and secure data storage. The chatbot will obtain explicit consent from users before collecting any personal data and will implement robust security protocols to protect sensitive information. Additionally, professional liability concerns will be addressed, especially in cases where inaccurate information might lead to legal risks. Ethical considerations, such as maintaining empathy and transparency in user interactions, are crucial to fostering trust between the technology and its users. CCeSca will also ensure accessibility for all, including individuals with disabilities, and comply with intellectual property laws to avoid copyright infringement. Through these measures, CCeSca will operate within the legal frameworks and promote responsible and inclusive technology usage.

References

[1] Business Times. (2024, January 31). Philippine economy posts South-East Asia's quickest expansion. Business Times. Retrieved from https://www.businesstimes.com.sg/international/asean/philippine-economy-posts-south-east-asias-quickest-expansion

[2] S&P Global Market Intelligence. (2023, December 22). Philippines economy shows robust growth into the new year. S&P Global Market Intelligence. Retrieved from https://www.spglobal.com/marketintelligence/en/mi/research-analysis/philippines-economy-shows-robust-growth-into-the-new-year.html

[3] Investopedia. (2023, June 28). Investment Fund. Investopedia. Retrieved from https://www.investopedia.com/terms/i/investment-fund.asp

[4] Lawrence, R. (1997). Using neural networks to forecast stock market prices. *University of Manitoba*, *333*(2006), 2013.

[5] Pang, X., Zhou, Y., Wang, P., Lin, W., & Chang, V. (2020). An innovative neural network approach for stock market prediction. *The Journal of Supercomputing*, *76*, 2098-2118.

[6] Egeli, B., Ozturan, M., & Badur, B. (2003). Stock market prediction using artificial neural networks. *Decision Support Systems*, *22*, 171-185.

[7] Indro, D. C., Jiang, C. X., Patuwo, B. E., & Zhang, G. P. (1999). Predicting mutual fund performance using artificial neural networks. *Omega*, *27*(3), 373-380.

[8] ÇEMREK, F., & DEMİR, Ö. (2021). Forecasting the returns of pension investment funds in Turkey with artificial neural network. *Cumhuriyet Science Journal*, *42*(4), 942-950.

[9] Wiliński, A., Smoliński, A., & Nowicki, W. (2016). Investment funds management strategy based on polynomial regression in machine learning. In *Intelligent Systems for Computer Modelling: Proceedings of the 1st European-Middle Asian Conference on Computer Modelling 2015, EMACOM 2015* (pp. 87-97). Springer International Publishing.

[10] Weiner, L. (2022). Comparing Actively Managed Mutual Fund Categories to Index Funds using Linear Regression Forecasting and Portfolio Optimization.

[11] Priyadarshini, E., & Babu, A. C. (2012). A comparative analysis for forecasting the NAV's of Indian mutual fund using multiple regression analysis and artificial neural networks. *International Journal of Trade, Economics and Finance*, *3*(5), 347.

[12] Chang, H. Y., Wen, C. H., & Pan, W. T. (2010). Prediction of the return of common fund through General Regression Neural Network. *Journal of Statistics and Management Systems*, *13*(3), 627-637.

[13] Sagir, A. M., & Sathasivam, S. (2017). The use of artificial neural network and multiple linear regressions for stock market forecasting. *Matematika*, *33*.

[14] Grossberg, S. (2013). Recurrent neural networks. *Scholarpedia*, *8*(2), 1888.

[15] Hope, T. M. (2020). Linear regression. In *Machine Learning* (pp. 67-81). Academic Press.

[16] Bwandowando. (n.d.). BPI Investment Funds Daily Data [Data file]. Kaggle. Retrieved from https://www.kaggle.com/datasets/bwandowando/bpi-investment-funds-daily-data

[17] a.aleksandrova@umni.bg. (2022, June 3). How to promote your Chatbot successfully | UMNI. UMNI. https://umni.bg/en/blog/how-to-promote-your-chatbot-successfully/: :text=Start%20promoting%20the%20chatbot%20by,or%20to%20create%20an%20avatar

[18] Necz, D. (2024). Rules over words: Regulation of chatbots in the legal market and ethical considerations. Hungarian Journal of Legal Studies, 64(3), 472–485. https://doi.org/10.1556/2052.2023.00472

[19] The General Data Protection Regulation. (2024, June 13). European Council Council of the European Union. Retrieved October 12, 2024, from https://europa.eu/!BnfjCg