Grenze International Journal of Engineering and Technology, January Issue

Recommendation System and Price Prediction using Deep Learning Model and Random Forest Regression

Simran Vedpathak[1], Seema Padghan[2], Sakshi Gaikwad[3] and Dr. Surekha Dholay[4]
[1]Sardar Patel Institute of Technology, Mumbai, Maharashtra, India
Email: simran.vedpathak@spit.ac.in
[2-4]Sardar Patel Institute of Technology, Mumbai, Maharashtra, India
Email: {seema.padghan, sakshi.gaikwad, surekha_dholay}@spit.ac.in

***Abstract*— This paper presents an integrated e-commerce framework using Convolutional Neural Networks (CNN) for query resolution and Random Forest Regression for price prediction. It employs web scraping to extract data from Amazon and Flipkart, generating relevant URL links based on user specifications, with MongoDB managing the dataset. Experimental results show the system's effectiveness in providing precise and tailored responses, enhancing the e-commerce experience.**

***Index Terms*– Deep Learning, Data Analysis Natural Language Processing, Chatbot, DataScraping, Python.**

I. Introduction

As technology continues to shape various facets of our lives, E-Commerce stands as a prominent product of this growth, particularly in the context of communication and messaging applications on smartphones. Chatbots bridges the gap between human and computer interaction, providing a unique avenue for Conversational Commerce. With the ability to comprehend context and deliver appropriate responses, chatbots facilitate commercial purposes through Conversational Commerce applications[10].

A. Study of ChatBots

The study compares ChatGPT, Google BARD, and Microsoft Bing, finding ChatGPT excels in accuracy, BARD in response time, and Bing in user satisfaction. For voice recognition in chatbots, smartphones with MIT App Inventor are recommended over the more complex Raspberry Pi [56]. AI integration in education, like ChatGPT, can transform teaching, but many teachers lack awareness and training. Proper support is crucial for teachers to ethically and effectively enhance learning with AI [7].

II. Literature Survey

A. Recommendation Systems in ChatBot:

Transitioning to urban navigation, the paper details the design and implementation of an inclusive chatbot-based interface within the CAPRIO system for pedestrian path recommendations, where text-based interface was built using both the Amazon Lex and MBF solutions to compare them and they found MBF to be more useful for text and voice-based interface[12].

Grenze ID: 01.GIJET.11.1.507_1

Authors in this paper [11] have developed a music recommendation system based on user emotion identified via chatbot. Utilized LSTM, Bidirectional LSTM, and 1-D CNN models trained on Twitter data. Bidirectional LSTM yielded best results for emotion classification into positive, negative, and neutral categories.
Utilizing a Generative Text Virtual Assistant with Recurrent Neural Network for Laptop Recommendation. Data gathering via UiPath precedes training and accessing laptop descriptions via

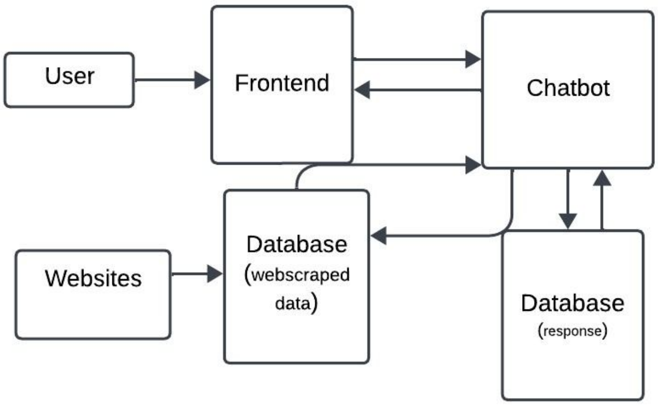

Figure 1. Architectural Flow Diagram

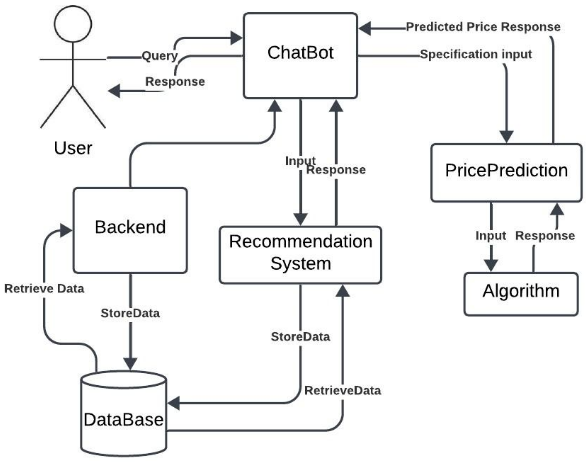

Figure 2. Overview of GuideMe Chatbot System for Recommendation and Price Prediction

website links. Achieved 96% accuracy using encoder-decoder method for predicting answers based on user questions [15].
Authors in this paper have implemented an intelligent travel chatbot on the Echo platform, gathering user preferences and modelling a collective knowledge base. Utilizes Restricted Boltzmann Machine (RBM) alongside Collaborative Filtering and DNN for recommendations. Integrates user data to enhance travel suggestions [13].
A Healthcare chatbot application [3] that identifies user symptoms through NLP, employing neural networks like MLP (ANN) for disease prediction and medication suggestion. While not a substitute for doctors, it offers valuable diagnoses, saving users time and money by reducing unnecessary doctor appointments.
This work [8] introduces a recommendation system utilizing chatbot technology and an auxiliary NLP library within a messaging application. Through an AHP-based comparison with a conventional ICD-10 application, the chatbot demonstrates effective ICD-10 code selection. Evaluation results endorse the chatbot’s superior performance and preference over traditional approaches. The study signifies a notable advancement in medical coding by seamlessly integrating chatbot technology with sophisticated decision-making processes. This research highlights the potential of chatbots in enhancing ICD-10 applications.

"DARA" [9] is integrated into a digital mental health cognitive bias intervention for anxiety. Developed using Juji, a chatbot platform supporting IDE-based and visual design. Design considerations include embodiment, response type, and conversational domain specificity.
Paper [1] introduces a recommendation system framework integrating chatbots for diverse scenarios. The backend engine computes user adherence profiles and offers proactive feedback via a chatbot frontend. Demonstrated feasibility with a Telegram chatbot managing hypertensive patients.

III. IMPLEMENTATION

The following section shows the outcomes of our entire project. The output is majorly focused in three ways- the proposed models the comparison of models and the application.

A. Proposed Models

Our chatbot algorithm, integrating Random Forest and Convolutional Neural Networks (CNN) models, was assessed using a publicly available dataset that included outcomes from various algorithms applied to the same data.

Background:
CNNs
In a typical CNN architecture for NLP tasks, such as text classification or sentiment analysis, the input data is typically represented as a sequence of word embeddings. The convolutional layer operates on these embeddings using filters to extract local patterns. Here's how a convolutional layer can be represented mathematically:

$$h_i = f\left(\sum_{j=1}^{k}(w_j * x_{i+j-1}) + b\right)$$

- h_i is the output feature map at position i.
- f is the activation function (e.g., ReLU).
- w_j is the filter/kernel of size k applied to the input sequence x starting at position i.
- b is the bias term.
- $*$ denotes the convolution operation.
- CNNs leverage weight sharing across the input space, allowing them to capture local patterns regardless of their position in the input sequence.

The comparison validated our chatbot's efficiency. Implemented in Python with scikit-learn, the Random Forest Regression achieved a training accuracy of 96% and a testing accuracy of 73% on an e-commerce dataset. These results showcase its robust learning and predictive abilities, highlighting effectiveness in capturing data patterns. The chatbot excels in delivering precise, budget-conscious suggestions for affordable laptops, affirming its capability in handling user preferences and laptop specifications.
The chatbot's development using deep CNN enhances its accuracy and introduces a valuable paradigm for natural language understanding and response generation.
The architecture captures intricate features in textual queries, improving responsiveness and handling nuanced language. This deep learning approach, with a specifically defined CNN architecture, excels in tasks requiring nuanced multiclass classification. Consequently, the chatbot stands out as an advanced conversational agent.

B. Comparison of Models

TABLE 1. PREDICTED PRICE OBTAINED FROM RANDOM FOREST REGRESSION (ACCURACY IS 72%)

Actual Value	Predicted Value	Difference
57990	63783.08	-5793.08
185790	75710.01	110079.99
61890	38227.19	23662.81
104990	109238.85	-4248.85
60990	71040.72	-10050.72
...	...	...
105990	78000.90	27989.10

Now, let's break down the differences between Random Forest Regression (RFR), Multiple Linear Regression (MLR), and Decision Tree Regression (DTR) in terms of equations:

Linear Regression (MLR)
The equation for multiple linear regression is:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon$$

- y is the dependent variable.
- $x_1, x_2, \dots, x_n$ are the independent variables.
- $\beta_0, \beta_1, \dots, \beta_n$ are the coefficients of the independent variables.
- ε is the error term.

MLR assumes a linear relationship between the independent variables and the dependent variable.

Decision Tree Regression (DTR)
Decision tree regression involves recursively partitioning the feature space into regions and fitting a simple model (typically constant) in each region. The prediction is the average (for regression) of the dependent variable in the training samples of the corresponding region. Let's represent a decision tree regression model as a piecewise constant function:

$$\hat{y} = \sum_{i=1}^{m} c_i \cdot I(x \in R_i)$$

- $\hat{y}$ is the predicted value.
- c_i is the constant predicted value in region R_i.
- $I(x \in R_i)$ is the indicator function that equals 1 if x belongs to region R_i, and 0 otherwise.
- m is the number of regions in the feature space.

Random Forest Regression (RFR)
Random forest regression combines multiple decision trees and averages their predictions. The prediction is the average (for regression) of the predictions of individual trees.

$$\hat{y} = \frac{1}{N_{trees}} \sum_{i=1}^{N_{trees}} \hat{y}_i$$

- $\hat{y}_i$ is the prediction of the i-th decision tree.
- N_{trees} is the number of trees in the random forest.

In summary, while MLR assumes a linear relationship between variables, DTR partitions the feature space and fits a simple model in each region, and RFR averages predictions from multiple decision trees, making it more flexible and capable of capturing complex relationships.

C. Application

In the implementation of Natural Language Processing (NLP), the NLTK library was judiciously employed, showcasing a comprehensive utilization of its features. The incorporation of speech recognition functionality within the system added a layer of complexity and user interaction, fostering a more dynamic and inclusive user experience. Notably, the NLTK library's specialized modules, such as the punkt library and WordNet, played pivotal roles in enhancing tokenization and semantic analysis. The punkt library facilitated precise sentence tokenization, contributing to the accurate parsing of input text, while WordNet enriched the system's lexical knowledge and semantic understanding. This thoughtful integration of NLTK's capabilities, including its support for speech recognition, tokenization, and lexical analysis, reflects a nuanced approach to NLP implementation, catering to diverse linguistic and interaction paradigms.

IV. DATASETS

For our Price Prediction model we have used a Dataset which contains 1300 entries with specification of ram, processor, ssd, size, weight and for recommendation we have a dataset which contains around 400 entries:
Price prediction dataset: {
"name":"Lenovo",
"processor":"Intel Core i5 11th Gen",

"price":"62990",
"ram":"16",
"OS":"Windows 11 Operating System",
"ssd":"512",
"size":"15.6" }
Recommendation System: {
"Brand": "ASUS",
"RAM": "16GB",
"SSD": "512GB SSD",
"Processor": "AMD Ryzen 5 5600H",
"Size": "15.6", "URL":
"https://www.amazon.in/dp/B0BMPPSLHZ", }

V. CONCLUSION

In conclusion, the developed chatbot has demonstrated remarkable capabilities in revolutionizing the laptop shopping experience. Leveraging a powerful combination of a Convolutional Neural Network (CNN) model for text understanding and a Random Forest Regression model for accurate price predictions, the chatbot excelled in recommending laptops from popular e-commerce platforms like Flipkart and Amazon. Its versatility was further enhanced by the incorporation of text-based and speech-recognition functionalities, providing users with a seamless and intuitive interaction.

The ability to comprehend natural language queries and predict prices accurately underscores the chatbot's effectiveness in assisting users with informed purchasing decisions.

FUTURE SCOPE

Looking ahead, future implementations could explore the integration of Natural Language Understanding (NLU) and advanced AI models to further elevate the chatbot's capabilities. By incorporating NLU, the chatbot could gain a deeper understanding of user intent and context, enabling more nuanced responses and refined recommendations.

Additionally, advanced AI models could enhance the chatbot's learning and adaptability, allowing it to stay updated with the evolving landscape of laptop specifications and user preferences. This integration of NLU and advanced AI would not only refine the chatbot's conversational abilities but also contribute to its continued growth as a sophisticated tool for personalized and intelligent assistance in the ever-expanding e-commerce domain.

REFERENCES

[1] S. Montagna, S. Mariani and M. F. Pengo, "A Chatbot-based Recommendation Framework for Hypertensive Patients," 2023 IEEE 36th International Symposium on Computer-Based Medical Systems (CBMS), L'Aquila, Italy, 2023, pp. 730-733, doi: 10.1109/CBMS58004.2023.00309.

[2] D. Van Thin, T. M. Quang, P. M. Toan, V. M. Thien, L. M. Hung and T. T. Quan, "A Humanlike Interactive Chatbot Framework for Vietnamese Banking Domain," 2022 9th NAFOSTED Conference on Information and Computer Science (NICS), Ho Chi Minh City, Vietnam, 2022, pp. 41-46, doi: 10.1109/NICS56915.2022.10013395.

[3] M. K. Ogirala, R. Tallapaneni, S. M. Chalamcharla and A. Chinta, "A Medical Diagnosis and Treatment Recommendation Chatbot using MLP," 2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, 2023, pp. 495-500, doi: 10.1109/ICAAIC56838.2023.10141211.

[4] S. Bhardwaz and J. Kumar, "An Extensive Comparative Analysis of Chatbot Technologies - ChatGPT, Google BARD and Microsoft Bing," 2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, 2023, pp. 673-679, doi: 10.1109/ICAAIC56838.2023.10140214.

[5] R. Sanjaya and C. Wibhowo, "Arduino, Raspberry Pi, and Smartphone Usage Comparison for Voice-based Virtual Assistant," 2022 6th International Conference on Information Technology (InCIT), Nonthaburi, Thailand, 2022, pp. 293-297, doi: 10.1109/InCIT56086.2022.10067741.

[6] E. R. Bekeš and V. Galzina, "Exploring the Pedagogical Use of AI-Powered Chatbots Educational Perceptions and Practices," 2023 46th MIPRO ICT and Electronics Convention (MIPRO), Opatija, Croatia, 2023, pp. 636-641, doi: 10.23919/MIPRO57284.2023.10159734.

[7] N. Siangchin and T. Samanchuen, "Chatbot Implementation for ICD-10 Recommendation System," 2019 International Conference on Engineering, Science, and Industrial Applications (ICESI), Tokyo, Japan, 2019, pp. 1-6, doi: 10.1109/ICESI.2019.8863009.

[8] R. X. Schwartz, A. Ramanan, D. Patel, A. Lynch, S. Baee and L. Barnes, "DARA: Development of a Chatbot Support System for an Anxiety Reduction Digital Intervention," 2022 Systems and Information Engineering Design Symposium (SIEDS), Charlottesville, VA, USA, 2022, pp. 139-144, doi: 10.1109/SIEDS55548.2022.9799419.

[9] M. Rakhra et al., "E-Commerce Assistance with a Smart Chatbot using Artificial Intelligence," 2021 2nd International Conference on Intelligent Engineering and Management (ICIEM), London, United Kingdom, 2021, pp. 144-148, doi: 10.1109/ICIEM51511.2021.9445316

[10] A. Nair, S. Pillai, G. S. Nair and A. T, "Emotion Based Music Playlist Recommendation System using Interactive Chatbot," 2021 6th International Conference on Communication and Electronics Systems (ICCES), Coimbatre, India, 2021, pp. 1767-1772, doi: 10.1109/ICCES51350.2021.9489138.

[11] B. T. Nixon, Z. Chen, S. Konduru, Y. Liu, C. Costa and P. K. Chrysanthis, "CAPRIO with Inclusive Pedestrian Path Recommendations," 2023 24th IEEE International Conference on Mobile Data Management (MDM), Singapore, Singapore, 2023, pp. 180-183, doi: 10.1109/MDM58254.2023.00037.

[12] A. Argal, S. Gupta, A. Modi, P. Pandey, S. Shim and C. Choo, "Intelligent travel chatbot for predictive recommendation in echo platform," 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 2018, pp. 176183, doi: 10.1109/CCWC.2018.8301732.

[13] N. Assavakamhaenghan, R. G. Kula and K. Matsumoto, "Interactive ChatBots for Software Engineering: A Case Study of Code Reviewer Recommendation," 2021 IEEE/ACIS 22nd International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), Taichung, Taiwan, 2021, pp. 262-266, doi: 10.1109/SNPD51163.2021.9704950.

[14] S. Tjayadi and V. C. Mawardi, "Laptop Recommendation Intelligent Virtual Assistant using Recurrent Neural Network with RPA for Data Scraping," 2022 IEEE 7th International Conference on Information Technology and Digital Applications (ICITDA), Yogyakarta, Indonesia, 2022, pp. 1-6, doi: 10.1109/ICITDA55840.2022.9971263.