engineering proceedings

MDPI

Proceeding Paper

Leveraging LLMs and Computer Vision for Personalized Nutrition Advice †

Mateo Tokić, Časlav Livada *, Tomislav Galba and Alfonzo Baumgartner

Faculty of Electrical Engineering, Computer Science and Information Technology Osijek, Josip Juraj Strossmayer University in Osijek, 31000 Osijek, Croatia; mateo.tokic@ferit.hr (M.T.); tomislav.galba@ferit.hr (T.G.); alfonzo.baumgartner@ferit.hr (A.B.)

* Correspondence: caslav.livada@ferit.hr

† Presented at the 34th International Scientific Conference on Organization and Technology of Maintenance (OTO 2025), Osijek, Croatia, 12 December 2025.

Abstract

This paper investigates the application of large language models (LLMs) in the domain of dietary advice, focusing on the recognition of ingredients and nutritional values from food products and the integration of this information into a system capable of delivering personalized recommendations. The research involved the development of a mobile application utilizing React Native and Python Flask frameworks. Optical character recognition (OCR) was implemented through the docTR model to extract nutritional information and ingredients from product images. Based on the extracted data and user profiles stored in a Firestore database, the system generates tailored dietary advice employing OpenAI's GPT-3.5-turbo model. The findings demonstrate the feasibility of using LLMs to provide personalized dietary recommendations, thereby opening new opportunities in the digital transformation of nutrition and dietary planning.

Keywords: large language models; dietary advice; optical character recognition; mobile applications; personalized nutrition

1. Introduction

In today's world, increasing attention is devoted to healthy lifestyle practices that include regular physical activity and proper nutrition. People of all age groups, constrained by modern lifestyles, experience reduced physical activity, with many spending the majority of their working hours in sedentary positions. Additionally, their diets are often based on high carbohydrate intake and fried foods containing excessive fats. With technological advancement, there has been a growing number of digital tools and applications that aim to help users track their dietary habits and optimize nutrient intake.

Traditional applications for dietary tracking often rely on simple databases or manual data entry, which can be tedious and error-prone. With the emergence of advanced technologies from the fields of machine learning and artificial intelligence, particularly Large Language Models (LLMs), opportunities arise for enhancing these applications through modern understanding and generation of personalized dietary advice [1].

LLMs, such as GPT (Generative Pre-trained Transformer) models, represent a contemporary approach to natural language processing, enabling computer systems to understand and generate text in a manner that mimics human communication. By applying these models in the nutrition domain, it becomes possible to develop applications that not only

Check for updates

Academic Editors: Mirko Karakašić, Naida Ademović, Tomislav Keser, Ercan Işık and Silva Lozančić

Published: 16 February 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

interpret nutritional information but also provide users with personalized advice based on their individual data, needs, preferences, and health goals [2].

This work introduces a novel approach that combines Optical Character Recognition (OCR) technology with LLM capabilities to automatically analyze images of nutritional values from food products and generate personalized meal recommendations. This technology represents a significant advancement over existing solutions by automating the process of nutritional data extraction and providing intelligent, contextualized dietary advice.

The main contribution of this research lies in the integration of OCR and LLM technologies to create a comprehensive system that can automatically recognize ingredients and nutritional values from product images, then generate personalized dietary recommendations tailored to individual user profiles. This approach significantly simplifies the decision-making process related to nutrition while providing a higher degree of personalization compared to existing market solutions.

The potential applications of this system extend beyond individual use to dietary counseling centers, healthcare institutions, and educational purposes for raising awareness about the importance of proper nutrition. Through further development and research in collaboration with nutritionists, this application could be adapted for real-world usage conditions and expanded to serve a broader user base.

2. Related Work

This section presents and compares several applications that provide dietary advice and enable users to track nutritional values and calorie intake throughout the day, based on user needs and goals. The application developed and presented in this work offers advice based on LLMs and differs from the following existing solutions.

2.1. Existing Nutritional Applications

OpenFoodFacts (Open Food Facts Association, Paris, France) is a collaborative database of food products where anyone can add and use its data. The mobile application's main feature is barcode scanning to display ingredients, nutritional values, and extensive product information with comparison capabilities. While users can set preferences for allergies, environmental impact, and dietary restrictions, the application primarily functions as a passive information display tool without providing personalized dietary recommendations [3].

MyFitnessPal (MyFitnessPal, Inc., San Francisco, CA, USA) offers daily calorie and nutritional intake recommendations based on user profiles and activity levels. The application provides various customizable plans tailored to different individuals and their needs. However, its main limitation is the reliance on manual data entry, which can be tedious and error-prone, especially for users tracking complex meals or unfamiliar products [4].

Noom (Noom, Inc., New York, NY, USA) is available for Android and iOS devices with a 14-day free trial followed by monthly subscription requirements. The application focuses on building healthy eating habits through daily lessons and provides progress tracking tools with personalized support via phone calls and in-app messaging. Despite its comprehensive approach, Noom's high cost and generalized advice limit its accessibility and personalization for individual nutritional needs [5].

PlateJoy (PlateJoy, LLC, San Francisco, CA, USA) uses questionnaires to collect details about users' lifestyle and cooking habits, then creates customized meal plans and grocery lists for households. The application is particularly useful for users who frequently cook at home and struggle with recipe variety or food waste management. However, it

https://doi.org/10.3390/engproc2026125021

focuses primarily on meal planning rather than real-time nutritional analysis and advice generation [6,7].

A comparison of the aforementioned applications is shown in Table 1.

Table 1. Overview of selected nutrition and diet-related applications, their primary functionalities, and main limitations.

Application	Main Functionality	Primary Limitation
OpenFoodFacts	Product information display	Passive data presentation only
MyFitnessPal	Calorie and nutrition tracking	Manual data entry required
Noom	Habit formation and coaching	Expensive, generalized advice
PlateJoy	Meal planning and shopping lists	Limited to planning, no real-time analysis

2.2. Positioning of the Proposed Approach

The existing applications demonstrate various approaches to nutritional guidance, from information databases to meal planning and habit formation. However, they share common limitations: passive information display, manual data entry requirements, high costs, or narrow functionality scope.

The proposed application addresses these gaps by combining Optical Character Recognition (OCR) technology with Large Language Models (LLMs) to automatically extract nutritional information from product images and generate personalized dietary advice. The proposed system employs OpenAI's GPT-3.5-turbo model, specifically the gpt-3.5-turbo-1106 version, accessed via the OpenAI API at the time of experimentation. This approach eliminates manual data entry while providing intelligent, contextualized recommendations that adapt to individual user profiles and real-time food choices. Unlike existing solutions that focus on single aspects of nutrition management, this system offers an integrated approach that bridges the gap between information acquisition and personalized advice generation.

3. Methodology

3.1. System Architecture

The nutritional mobile application was developed using a modern full-stack architecture that ensures an intuitive user experience and secure, high-quality data processing. User authentication, management, and updating of personal data relevant to nutrition advice, as well as the ability to obtain meal and product recommendations via images of ingredients and nutrition tables, are core requirements. The client side of the application, built with React Native, provides a cross-platform mobile interface (for both iOS and Android) using JavaScript and JSX syntax for declarative and maintainable UI development [8]. The frontend stack consisted of React Native v0.72, running on Node.js v18, with source code written in JavaScript (ECMAScript 2021) and JSX syntax compiled by the React v18 toolchain. React Native's integration with the React library enables code reuse, interaction with native components, and access to a large ecosystem of pre-built solutions and documentation, ensuring scalability and fast development cycles as can be seen on Figure 1.

On the server side, Python 3.11 and the Flask 2.3 web framework Flask is used as a minimalist web framework tailored for small-to-medium applications and secure REST API services [9]. Flask's flexibility allows for customized project structuring and the addition of application logic as required, while the defined API endpoints standardize communication between the frontend and backend. For managing authentication and real-time data, the application leverages Firebase. Firebase Authentication provides secure user sign-in through email, password, phone, or social identity providers (Google, Facebook, Twitter), adhering

to industry standards like OAuth 2.0 and OpenID Connect. Real-time storage and retrieval of user profiles are handled by Cloud Firestore [10], enabling dynamic updates and synchronization across user devices. By connecting these components, the architecture allows mobile users to upload images of ingredient lists and nutrition tables, which are processed through backend machine learning models for OCR and LLM-powered advice generation. The result is a scalable, reliable, and user-friendly mobile application architecture capable of fulfilling the requirements for personalized, AI-driven dietary guidance.

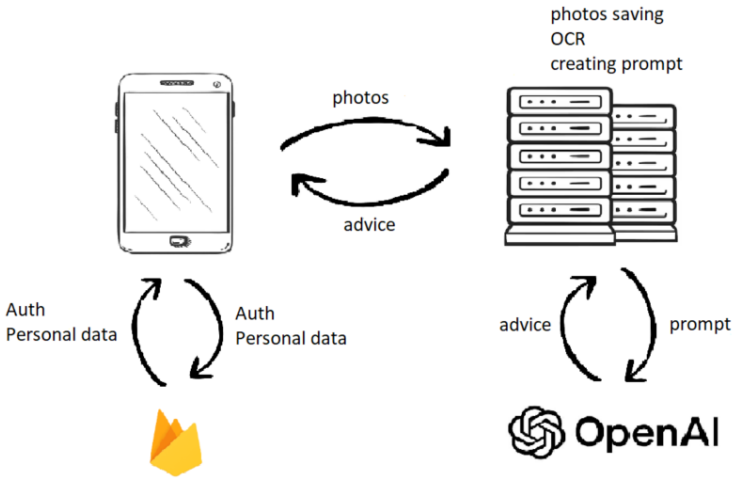


Figure 1. Architecture of the application.

3.2. OCR Implementation

Optical Character Recognition (OCR) is a technology that converts images containing text into machine-readable data, enabling digital processing and analysis of printed or handwritten information. In this application, OCR is essential for extracting ingredient lists and nutritional values directly from photographs of food product packaging, which helps automate and simplify the user experience by removing the need for manual data entry.

The system uses the DocTR (Document Image Transformer) model, an advanced, transformer-based OCR solution optimized for handling common challenges such as geometric distortions and inconsistent lighting often present in user-captured images. DocTR consists of two main modules: a Geometric Unwarping transformer that corrects misalignments and warps in images, and an Illumination Correction transformer that adjusts lighting variations. Utilizing self-attention mechanisms, the model accurately reconstructs sequences of characters, words, and sentences even from suboptimal image inputs [11].

When a user uploads photographs of a product's ingredients and nutrition table, the backend service processes these images with the DocTR model, yielding structured text data. Custom functions then parse this output to extract relevant information. The *extract_ingredients* function scans the OCR text for keywords such as "ingredients" to build a structured ingredient list, robustly handling multi-word items and punctuation. The *extract_nutrition_values* function leverages regular expressions and keyword mappings to identify and collect all key nutritional metrics, ensuring numerical accuracy and recognizable value-unit pairs (e.g., "12.4 g", "230 kcal").

The extracted data forms the foundation for generating personalized nutrition advice. The application combines user profile details with OCR-extracted results, forwarding them to a language model that creates relevant dietary recommendations. By automating the entire pathway from image capture to structured data extraction, this OCR implementation enables user-friendly, scalable and reliable integration of real-world product information into digital personalized nutrition management.

3.3. LLM Integration

For personalized dietary advice generation, the application integrates OpenAI's GPT-3.5-turbo, a leading model from the GPT-3.5 series developed in 2022. Chosen for its speed, cost-effectiveness, and sufficiently strong performance for most needs, GPT-3.5-turbo remains in active use despite the release of newer GPT-4 models. This model is trained on a dataset updated until September 2021 and is fine-tuned via reinforcement learning with human feedback (RLHF). RLHF enables the model to better align with human preferences by incorporating feedback directly from human evaluators during training, thus improving response relevance and safety compared to models relying solely on supervised or unsupervised learning [12].

One notable challenge when using large language models like GPT-3.5 is hallucination, the phenomenon where the model may provide factually incorrect or fabricated responses. These issues can stem from ambiguous prompts, outdated training data, or the model's inability to access current information.

In this application, a dedicated function generates personalized prompts using user profile data, extracted nutritional values, and ingredient lists. The prompt is formatted as a concise query with instructions for the model to act as a nutrition expert, to provide brief, relevant advice, recommend serving sizes, and suggest additional ingredients if appropriate [13].

The prompt is sent to the OpenAI API using the gpt-3.5-turbo endpoint, with the role of the model explicitly set as a nutritionist. The length of generated responses is limited to 250 tokens to ensure brevity and clarity, while the temperature parameter is set to 0.4 to promote consistency and reduce hallucination risk. The function returns the first API response containing the model's personalized dietary recommendation, thus tightly integrating LLM capabilities into the user advice workflow [14].

3.4. Data Processing Pipeline

The data processing pipeline transforms a user's photo of a food label into a personalized nutritional recommendation in several automated steps. First, the user captures or uploads an image of the food product's ingredient list or nutrition facts using the mobile application. The image is preprocessed through steps like resizing, contrast adjustment, binarization, and noise reduction to enhance clarity and maximize OCR accuracy. The preprocessed image is then sent to an OCR engine, which detects and recognizes text areas, converting the visual content into machine-readable format.

Next, custom parsing functions analyze the extracted text: keywords signal the start of ingredient lists, and regular expressions or mappings are used to identify nutritional values like "energy," "protein," or "fat." Detected data is structured into standardized JSON, allowing for reliable downstream processing. This structured information—combined with the user's profile data—is then used to build a prompt for a Large Language Model (LLM). The user's details, ingredient list, and nutrition data form the context for the LLM, which generates concise, relevant dietary advice tailored to the user's preferences and health goals.

Finally, the application displays this personalized recommendation, completing an end-to-end pipeline: from mobile photo capture, through intelligent image and text processing, to actionable nutrition guidance—automated, reliable, and user-friendly.

4. Results and Evaluation

In this chapter, examples of the application's functionality will be presented using the technologies mentioned above. After taking a photo of the ingredients by tapping the camera icon, the user is prompted to photograph the nutritional information table as well.

For the purpose of demonstrating how the system generates dietary recommendations, a packaged ham product was selected. Its list of ingredients and nutritional values are shown in Figure 2. These visual inputs are processed by the application to provide personalized suggestions based on the user's dietary needs and preferences. For understanding the Figure is translated – Average nutritional values per 100 g of product: Energy 410 kJ/98 kcal, Fat 3.6 g, of which saturated fatty acids 1.4 g, Carbohydrates 3.3 g, of which sugars 1.0 g, Protein 13 g, Salt 2.4 g.

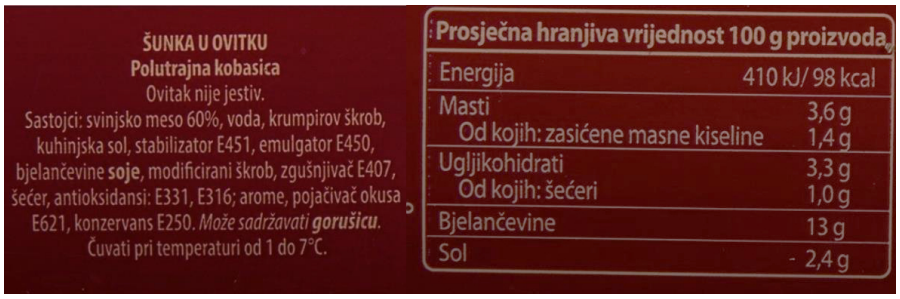


Figure 2. Ingredients of the ham.

By photographing this product, JSON documents are automatically generated on the server side, containing the values that the docTR model successfully extracted from the images. These values are the result of optical character recognition (OCR) applied to both the list of ingredients and the nutritional information table.

Figure 3 shows the recognized ingredients and nutritional values, which are saved in two separate files: *ingredients.json* and *nutritional_values.json*, located within the product directory. These structured data files serve as input for further processing, such as nutritional analysis or the generation of personalized dietary recommendations.

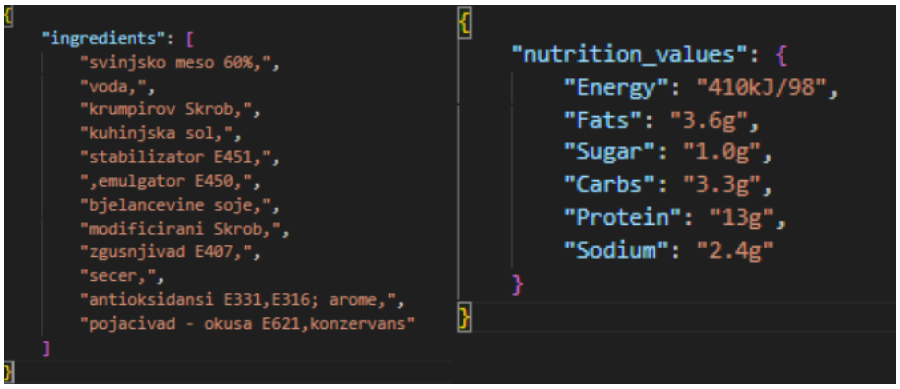


Figure 3. Example of JSON file.

The extracted data about the food product, along with the user's personal information, were used to generate a personalized recommendation, as shown in Figure 4.

This advice is tailored to the user's dietary profile and preferences, combining product-specific nutritional data with user-specific health goals or restrictions, thereby demonstrating the practical application of intelligent recommendation systems.

https://doi.org/10.3390/engproc2026125021

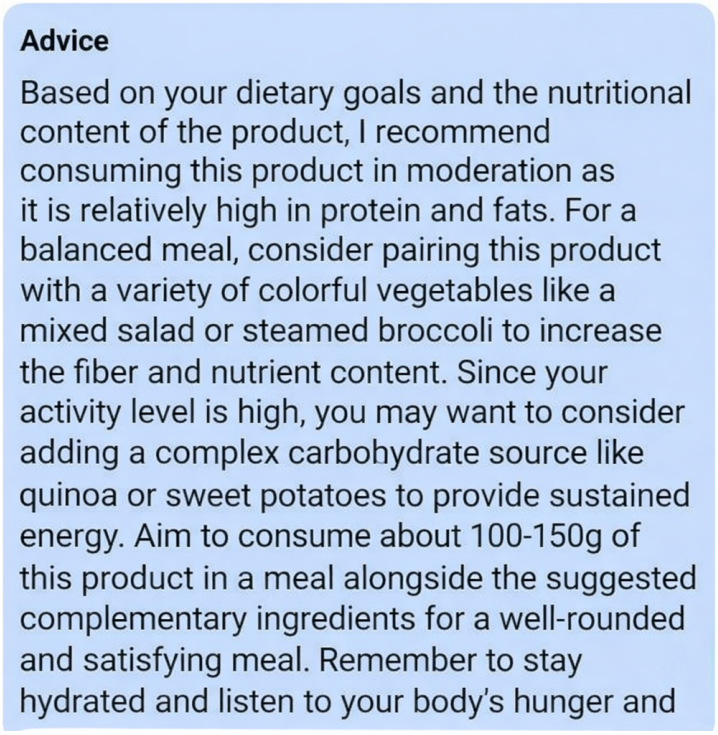


Figure 4. Recommendation regarding a food product.

5. Discussion

While the application fulfilled the core objectives of the paper, several limitations were observed that highlight opportunities for future improvement. One of the key challenges lies in the image processing and OCR (Optical Character Recognition) phase. Text recognition accuracy is highly dependent on image quality, background color, and inconsistent lighting conditions. These factors significantly affect the system's ability to extract reliable data from product labels.

An additional issue relates to the recognition of characters specific to the Croatian alphabet. Letters containing diacritical marks—such as č, ć, đ, š, and ž—were often misinterpreted or omitted, reducing the reliability of the parsed data.

The quality of the generated recommendations is also constrained by the limited set of user input parameters used in this version of the application. Personalized dietary guidance would benefit from incorporating a broader range of user data, including age, gender, activity level, body measurements, eating habits, and health conditions.

In future work, the accuracy of OCR for the Croatian language will be improved through fine-tuning on a dedicated dataset of local food products, and parsing functions will be enhanced to better handle diverse label formats and layouts. User profiling will be expanded to include parameters such as age, gender, physical activity level, dietary restrictions, and specific health goals to enable more personalized nutritional advice. The GPT-3.5-turbo model will be fine-tuned on a domain-specific dataset curated by nutrition professionals, ensuring higher quality, reliability, and scientific validity of the recommendations. Additionally, integration with wearable devices will be implemented to allow automatic synchronization of activity and caloric data, resulting in more dynamic and context-aware dietary recommendations.

6. Conclusions

The development of the mobile application successfully met the objectives outlined in this paper. By utilizing the React Native framework with JavaScript, a functional and secure

client-side application was created, featuring a user-friendly interface and the ability to capture images, all while ensuring reliable communication with the server component. The server side, built using the Flask framework in Python, efficiently handled data exchange with the client, demonstrating stability and responsiveness.

Integration of Firebase Authentication and Firestore Database proved to be an effective solution for managing user registration, login, and the secure storage and modification of personal data. These technologies not only simplified the development process but also contributed to the application's scalability and overall security.

Given its capabilities, the application has potential for real-world use in healthcare settings, where it could assist in monitoring patient nutrition and supporting dietary planning. Ultimately, this work demonstrates the feasibility of partially automating nutritional guidance through modern mobile and AI technologies, offering a foundation for more advanced and personalized health-support tools in the future.

Author Contributions: Conceptualization, M.T. and Č.L.; methodology, Č.L.; software, T.G. and A.B.; validation, M.T., Č.L. and T.G.; formal analysis, A.B.; investigation, M.T.; resources, M.T.; data curation, M.T. and T.G.; writing—original draft preparation, Č.L.; writing—review and editing, Č.L., T.G. and A.B.; visualization, M.T.; supervision, Č.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Solomon, T.P.J.; Laye, M.J. The sports nutrition knowledge of large language model (LLM) artificial intelligence (AI) chatbots: An assessment of accuracy, completeness, clarity, quality of evidence, and test-retest reliability. *PLoS ONE* **2025**, *20*, e0325982. [CrossRef] [PubMed]
2. Gençer Bingöl, F.; Ağagündüz, D.; Bingol, M.C. Accuracy of Current Large Language Models and the Retrieval-Augmented Generation Model in Determining Dietary Principles in Chronic Kidney Disease. *J. Ren. Nutr.* **2025**, *35*, 401–409. [CrossRef] [PubMed]
3. OpenFoodFacts. 2024. Available online: https://world.openfoodfacts.org/discover (accessed on 19 November 2024).
4. Morell, M.; DeSoto, L.; Kacyon, C.; Robinson, L. 7 Best Nutrition Apps: Tried and Tested in 2024. 2024. Available online: https://www.helpguide.org/handbook/healthy-living/best-nutrition-apps (accessed on 19 November 2024).
5. Forman, S. The Noom Trial 2024: Here's Everything You Need to Know to Get Started. 2022. Available online: https://www.usmagazine.com/shop-withus/news/noom-free-trial-exercise-diet-weight-loss-health-home/ (accessed on 19 November 2024).
6. Ajmera, R. A Dietitian's Picks of the 10 Best Nutrition Apps to Download. 2022. Available online: https://www.healthline.com/nutrition/top-iphone-android-apps (accessed on 19 November 2024).
7. Strong, L. PlateJoy Review (2024): Meal Planning Just Got Easier with This App. 2024. Available online: https://www.garagegymreviews.com/platejoy-review (accessed on 19 November 2024).
8. React. Writing Markup with JSX. 2024. Available online: https://react.dev/learn/writing-markup-with-jsx (accessed on 22 November 2024).
9. Grinberg, M. *Flask Web Development*; O'Reilly Media: Sebastopol, CA, USA, 2014.
10. Google Firebase. 2024. Available online: https://firebase.google.com/docs/firestore (accessed on 8 November 2024).
11. IBM. What Is Optical Character Recognition (OCR)? 2024. Available online: https://www.ibm.com/think/topics/optical-character-recognition (accessed on 8 November 2024).
12. Ipsen, A. GPT Base, GPT-3.5 Turbo & GPT-4: What's the Difference? 2023. Available online: https://www.pluralsight.com/resources/blog/ai-and-data/ai-gpt-models-differences (accessed on 7 November 2024).

13. Mandour, A. GPT-3.5 Model Architecture. 2024. Available online: https://iq.opengenus.org/gpt-3-5-model/ (accessed on 19 November 2024).
14. Mittal, A. Tackling Hallucination in Large Language Models: A Survey of Cutting-Edge Techniques. 2024. Available online: https://www.unite.ai/tackling-hallucination-in-large-language-models-a-survey-of-cutting-edge-techniques/ (accessed on 17 November 2024).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.