International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 11 | Nov - 2026 SJIF Rating: 8.659 ISSN: 2582-3930

Health LLM: Transformer-Driven Personalized Wellness Companion for Children

Ankolu Chandra Sekhar[1], Satharla Siddartha[2]

Department of Artificial Intelligence and Data Science
Indian Institute of Information Technology Design and Manufacturing, Kurnool
Email: 226cs2013@iiitk.ac.in

Abstract: Children's wellness is influenced by interconnected factors such as nutrition, sleep, physical activity, screen exposure and emotional well-being. Conventional tracking applications commonly present isolated measurements or generic advice, making it difficult for parents to interpret several lifestyle indicators together. This paper presents HEALTH LLM, a transformer-driven wellness companion that converts a child-specific profile into structured, understandable wellness guidance. The application is implemented using Python and Streamlit. A Pydantic data model validates profile information, including age, gender, height, weight, sleep duration, screen time, activity level, eating habits, mood and reported symptoms. The validated profile is transformed into a structured prompt and submitted to OpenAI's GPT-4o-mini model through the OpenAI API. The response is requested in a structured JSON format and presented as a wellness score, daily plan, nutrition advice, physical activity suggestions, mental-wellness support, risk assessment and parent/guardian guidance. The project report documents functional operation using sample child profiles and demonstrates an end-to-end workflow from data entry to recommendation display. The system is intended for preventive wellness education and decision support; it is not a diagnostic tool and has not been clinically validated. The paper describes its architecture, implementation, observed functionality, limitations and directions for responsible extension.

Keywords: Child wellness, Large Language Models, Transformer, Personalized healthcare, GPT-4o-mini, Streamlit.

I. INTRODUCTION

Childhood is a period of rapid physical, cognitive and emotional development. Everyday behaviours—including food choices, sleep routines, physical activity, recreational screen use and emotional experiences—contribute to a child's overall well-being. Parents and guardians are often required to interpret these factors together, although the information may be scattered across routine observations, growth records and separate tracking applications. Tools that monitor only one parameter may not explain how different lifestyle patterns relate to one another.

Artificial intelligence (AI) has created opportunities to process diverse information and provide accessible digital assistance. Natural language processing (NLP) systems can interpret user-provided descriptions, while Large Language Models (LLMs) can generate context-sensitive text. The Transformer architecture introduced an attention-based approach to sequence processing and became a foundation for many contemporary language models [1]. Generative pre-trained models subsequently demonstrated the ability to perform a range of language tasks through prompting [2]. These capabilities motivate the development of interactive wellness applications that translate structured user information into readable guidance.

This paper presents HEALTH LLM: Transformer-Driven Personalized Wellness Companion for Children. The system collects a child-specific profile through a web interface, validates the information, constructs a structured prompt and uses GPT-4o-mini to generate a wellness report. The report organizes its output into multiple categories so that parents can review lifestyle-related suggestions in one place.

The main contributions of this work are: (1) a unified profile representation covering physical, behavioural and emotional wellness inputs; (2) a validation and prompt-generation pipeline that connects structured data with an LLM; (3) a structured output format designed for parent-friendly presentation; and (4) a functional web application demonstrating the end-to-end workflow. The work is an implementation study rather than a clinical efficacy study. No claim of diagnostic accuracy or measured improvement in child health is made.

II. LITERATURE REVIEW

Miotto et al. reviewed deep learning applications in healthcare and discussed their potential for extracting patterns from complex health information, alongside challenges such as interpretability, data availability and privacy [3]. Their review provides a broader foundation for applying AI to health-related information, while also highlighting the importance of responsible deployment.

Brown et al. demonstrated that large autoregressive language models can perform multiple language tasks through few-shot prompting [2]. This prompted the wider use of general-purpose LLMs in interactive applications. Vaswani et al. introduced the Transformer, whose self-attention mechanism supports contextual processing of sequences without recurrent computation [1]. These developments underpin the language-generation capability used in HEALTH LLM.

In healthcare, Singhal et al. evaluated Med-PaLM 2 on medical question-answering benchmarks and included expert and layperson assessments of answer quality and safety [4]. Their work illustrates both the potential of LLMs in health-related communication and the need for domain-specific evaluation. Sallam's review of generative AI in healthcare also discusses opportunities in communication and support, while identifying concerns involving reliability, privacy and ethical use [5].

The reviewed literature establishes the technical potential of transformers and LLMs, but does not by itself validate the specific recommendations produced by the present application. HEALTH LLM focuses on a narrower software-engineering contribution: combining a structured child wellness profile, input validation, prompt construction and categorized output presentation in a lightweight web application.

A. Research Gap

Many basic wellness tools emphasize isolated tracking or fixed recommendations. A conversational LLM can produce more flexible explanations, but unstructured input and output can make application behaviour difficult to validate and display consistently. The project addresses this software-level gap by introducing a typed child profile, validation before model invocation and a defined JSON-oriented response structure. The current implementation does not close the broader clinical evidence gap; pediatric expert review and systematic safety testing remain necessary.

III. SYSTEM DESIGN AND METHODOLOGY

A. Proposed System

HEALTH LLM is designed as a parent-facing wellness assistance application. The parent or guardian enters the child's details using a Streamlit interface. The application validates the submitted values through a Pydantic model named ChildHealthProfile. Validated information is passed to a prompt-generation component, which prepares instructions and profile context for GPT-4o-mini. The model response is processed and displayed through the dashboard.

The system is modular. Its principal components are the user interface, data validation, configuration management, prompt builder, health analysis engine, result processing and recommendation dashboard. This separation makes individual functions easier to maintain and provides clear boundaries between user input, external model communication and presentation.

B. Input Collection and Validation

The interface collects the child's name, age, gender, height, weight, sleep duration, screen time, physical activity level, eating habits, mood and symptoms. These inputs provide a combined view of selected lifestyle and wellness factors. Pydantic is used to define the expected profile structure and validate data types, required fields and configured value constraints. Invalid or incomplete submissions are rejected before the model request, and the interface can communicate the validation issue to the user.

The validation stage improves structural consistency; it does not establish that a reported value is medically accurate. The system depends on the parent or guardian to provide correct information. The project report does not describe a separate pediatric dataset or sensor-based measurement pipeline. In a practical deployment, the input form should explain each field, specify the unit expected for measurements and distinguish optional observations from required profile attributes. For example, height should be recorded in centimetres, weight in kilograms, and sleep and screen exposure as durations. Consistent units reduce ambiguity when the profile is transformed into prompt context. The

interface should also allow a guardian to review the entered profile before submitting it to the external model service.

The collected variables represent different dimensions of wellness. Age and gender provide basic context; height and weight describe body measurements; sleep duration and screen time represent selected daily routines; activity level and eating habits capture behavioural factors; and mood and symptoms allow the guardian to communicate observations that cannot be represented by numeric measurements alone. The combination is intended to give the language model a broader context than a single measurement. Nevertheless, these fields are not a complete pediatric assessment. They omit many clinical, developmental, family and environmental factors, and the application should not infer a medical condition from their presence or absence.

C. Prompt Construction and Model Processing

After validation, the profile is converted into a structured representation. The prompt builder combines the profile fields with instructions specifying the expected categories of the wellness response. The Health Analysis Engine sends this prompt to GPT-4o-mini through the OpenAI API. The model is requested to return structured JSON so that the application can identify and display the expected fields.

The requested response categories are: wellness score, daily wellness plan, nutrition advice, physical activity recommendations, mental-health support, risk assessment and parent/guardian guidance. The wellness score is a model-generated summary within the application; the project report does not establish it as a clinically validated scale. Similarly, the risk assessment is informational output and must not be interpreted as a diagnosis. The use of a fixed response schema serves two purposes. First, it gives the model a clear output contract and reduces the likelihood that an expected section is omitted. Second, it allows the application to map returned fields to corresponding dashboard components without requiring the user to interpret a long, unstructured paragraph. If a response is malformed or a field is absent, a production implementation should detect the condition and present a safe fallback rather than treating incomplete output as a complete assessment.

Prompt construction is an important intermediary in this design. A profile is not simply forwarded as an isolated string; it is combined with instructions that define the intended role of the application, the audience for the response and the categories expected in the result. The prompt should encourage age-aware, practical and non-alarmist language, request that uncertainty be acknowledged, and avoid definitive diagnosis or medication instructions. Prompt instructions can influence the style and organization of a response, but they cannot guarantee factual correctness. The application therefore treats model output as generated guidance that requires cautious interpretation rather than as verified clinical knowledge.

D. Result Processing and Presentation

The response-processing module extracts the returned content and organizes it into readable sections. The Streamlit dashboard presents the wellness score using metric components and displays the recommendations under separate headings. This design aims to make the generated information easier for a parent or guardian to review and act upon. The application flow is summarized in Figure 1.

The dashboard separates information into categories that correspond to everyday decisions. Nutrition advice can be reviewed alongside eating habits; activity suggestions can be considered in relation to the reported activity level; and sleep or screen-related recommendations can be discussed as part of a daily routine. Mental-wellness support and parent/guardian guidance are included to avoid limiting the report to physical measurements alone. The categories are intended as a readable organization of model-generated content and should not be interpreted as independent clinical assessments. A guardian remains responsible for deciding whether a suggestion is appropriate for the child's circumstances and seeking professional advice where required.

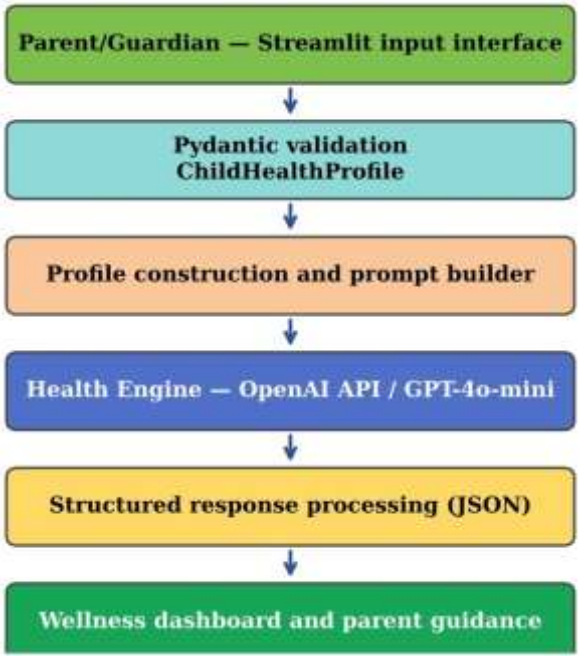


Figure 1 : ARCHITECTURE AND INFORMATION FLOW OF HEALTH LLM

IV. IMPLEMENTATION

A. Technology Stack

Python is used as the primary programming language. Streamlit provides the interactive web interface, while Pydantic defines and validates the child health profile. The OpenAI Python SDK provides access to GPT-4o-mini through the OpenAI API. The python-dotenv package loads configuration values from environment variables, and JSON is used as the structured response representation. The technologies are listed in Table 1.

Table 1 : IMPLEMENTATION TECHNOLOGIES

Technology	Role in the application
Python	Core application logic
Streamlit	Web interface and dashboard
Pydantic	Profile schema and input validation
OpenAI API	Communication with the language model
GPT-4o-mini	Generation of wellness guidance
python-dotenv	Environment-based configuration
JSON	Structured model response handling

B. Application Modules

User interface: Streamlit input controls collect the child's profile and provide the action to generate a wellness report. The layout is designed for parents and guardians and does not require them to interact directly with the API.

Data validation: The ChildHealthProfile model checks the structure of the submitted information before further processing. This helps prevent malformed values from being passed to the prompt builder.

Configuration: API credentials and model settings are loaded through environment variables using a .env configuration file and python-dotenv. Credentials should not be embedded in publicly shared source code.

Prompt builder and Health Engine: The prompt builder converts the validated profile into model-readable context. The Health Engine uses the OpenAI SDK to send the request and receive the generated response from GPT-4o-mini.

Result processing and dashboard: The returned JSON content is extracted and arranged into the wellness categories.

The dashboard displays the generated score and explanatory recommendations in separate sections.

C. Operational Workflow

The operational sequence begins when a parent or guardian opens the application and enters the child's information. The profile is validated. If validation fails, the request should not proceed to the model. If validation succeeds, the profile is formatted into a prompt and submitted through the API. The response is parsed, organized and rendered on the dashboard. The resulting recommendations can help users consider routine lifestyle changes, while any medical concern must be discussed with a qualified healthcare professional.

From a software perspective, the workflow can be considered as a sequence of controlled transformations: raw form values are converted into a validated profile; the profile is converted into prompt context; the external model response is converted into structured application data; and the structured data is converted into interface components. Each transformation creates a point at which errors can be detected. Input validation addresses errors at the first boundary, response parsing addresses formatting problems at the model boundary, and presentation logic determines how missing or unexpected fields are communicated. This modular approach is useful because failures in an external model service should not be confused with errors in the user-entered profile or the dashboard itself.

The configuration module also has an operational role. The API key is loaded from environment configuration rather than being placed directly in the application source. This reduces the risk of accidentally exposing credentials when source files are shared. In a deployed version, the key should be stored in a protected secrets manager, request access should be restricted, and logging should avoid recording child-specific profile information. The report describes the use of a local .env file and python-dotenv; this is a development configuration approach and should be strengthened before public deployment.

D. Processing Logic and Design Boundaries

The principal processing logic can be summarized as a guarded pipeline. First, the application receives the values entered in the form. Second, it constructs a ChildHealthProfile object and checks whether the values conform to the declared schema. Third, the validated object is serialized into prompt context and combined with the response instructions. Fourth, the Health Engine submits the request to the configured model endpoint. Fifth, the returned content is checked and organized for display. Finally, the dashboard renders the available recommendation categories. The validation gate is important: invalid profile data should not be submitted as if it were a complete record.

The application has a clear boundary between deterministic software operations and probabilistic model generation. Form handling, schema validation, configuration loading and dashboard rendering are conventional application functions. The natural-language analysis and recommendation text are generated by the external LLM and may not be identical for repeated requests. This distinction affects testing. Deterministic modules can be checked against expected inputs and outputs, whereas generated content requires qualitative review against defined criteria. A production implementation should retain this separation in its logging and error-handling strategy.

The project report describes the use of a structured JSON response, but does not specify a complete schema-validation library for the model output. A future implementation should validate the response independently before rendering it. Required fields should be checked, unexpected data types should be rejected or safely normalized, and unavailable sections should be clearly identified. The application should also handle network errors, API rate limits, empty responses and service timeouts without displaying partial or misleading guidance as a complete wellness report. The documented behaviour at each stage is summarized in Table 2.

Table 2 : FUNCTIONAL OBSERVATIONS REPORTED FOR THE PROTOTYPE

Stage	Documented behaviour
Profile entry	Collects demographic, body-measurement and lifestyle inputs.
Validation	Uses Pydantic to check required fields, data types and configured constraints.
AI request	Sends the validated profile as structured prompt context to GPT-4o-mini.
Response	Requests wellness guidance in categorized, JSON-oriented form.

Dashboard	Presents wellness score and nutrition, activity, mental-wellness and parent guidance sections.

The intended user is a parent or guardian rather than a healthcare professional. Accordingly, the interface should use plain language and make the limitations of the generated content visible near the results. The system should not provide medication dosage, diagnose disease or encourage users to delay professional assessment. When a guardian enters a potentially urgent symptom, the safer design is to show a clear instruction to contact an appropriate healthcare provider or emergency service instead of relying on generated recommendations. These safeguards are design requirements for responsible future deployment and are not claimed as fully implemented features in the current prototype.

V. RESULTS AND DISCUSSION

The project report documents the working application and its use with sample child health profiles. The reported demonstration covers profile entry, validation, model processing and presentation of a generated wellness report. The application home screen provides the input interface; a completed profile screen shows sample information; and the generated report presents the model output in categorized sections.

The documented sample report identifies lifestyle-related concerns such as low physical activity, unhealthy dietary patterns, excessive screen exposure and inadequate sleep, and provides corresponding suggestions. Examples described in the report include increasing vegetable intake and hydration, choosing balanced meals, replacing less nutritious snacks, encouraging outdoor activity, establishing consistent sleep routines and monitoring screen use. These are examples of generated guidance, not independently verified clinical findings.

The implementation demonstrates the feasibility of integrating a typed input model with a general-purpose LLM in a lightweight web application. Pydantic provides a boundary between raw user input and downstream processing, while structured prompting and response organization support a predictable presentation layer. Streamlit reduces the effort required to expose the workflow through a browser-based interface.

The three application views described in the project report represent complementary stages of use. The home screen provides the entry point and collects the child's profile. The profile input screen demonstrates a completed sample record, illustrating the range of information supplied to the application. The generated personalized health report demonstrates the final presentation stage, where the output is arranged into a summary and recommendation categories. Together, these views show that the prototype is not limited to an isolated model call; it connects data collection, validation, inference and display in a single user-facing workflow.

The generated report described in the project includes a summary of factors that may affect the child's wellness, followed by practical suggestions. Nutrition suggestions include balanced meals, greater vegetable intake, hydration and healthier snack alternatives. Physical activity suggestions include sports, outdoor games and regular movement. Sleep-related guidance includes maintaining a consistent schedule and bedtime routine. Screen-use guidance encourages monitoring and limits as appropriate. These examples illustrate the type of output produced by the prototype. They are not prescriptions, and their suitability depends on the child's age, health status, culture, preferences and individual circumstances.

The choice of a general-purpose model makes the prototype comparatively easy to develop because it does not require the project team to train a language model from scratch. At the same time, it creates dependencies on an external API, model availability, service latency, usage cost and the provider's data-handling terms. The response may also change as the underlying model or service is updated. Reproducible evaluation should therefore record the model identifier, prompt version, relevant configuration and date of each test, while avoiding the storage of unnecessary personal information.

However, the evaluation described in the college report is primarily functional and illustrative. It does not provide a defined test-set size, a validated pediatric benchmark, clinician-rated recommendation quality, response-latency measurements, cost analysis, or quantitative accuracy metrics. Consequently, numerical performance claims cannot be supported by the available project evidence. The system's outputs may vary with model behaviour and prompt changes, and generated language can contain omissions or inaccurate statements. A structured JSON response improves parsing but does not guarantee medical correctness.

Because the application concerns children, privacy and safety require particular attention. Only necessary information should be collected, access to API credentials must be protected, and personal data should not be retained or shared beyond the intended purpose. The interface should clearly communicate that the system provides general wellness education, not diagnosis, treatment or emergency advice. High-risk symptoms or urgent concerns should be directed to professional care rather than handled solely by an LLM.

A more complete evaluation should examine the application at several levels. Functional testing should verify that valid profiles can be submitted, invalid profiles are rejected, API failures are handled gracefully, and the dashboard displays all expected categories. Output-quality evaluation should involve qualified pediatric professionals who assess factual correctness, age appropriateness, completeness, clarity and potential harm. Usability testing should determine whether parents understand the wellness score, recognize the limitations of generated guidance and can locate the appropriate recommendation sections. Technical evaluation should record response time, API failure rate and resource or service costs under a clearly specified test protocol. None of these measurements is reported quantitatively in the current college report, so they are proposed evaluation dimensions rather than completed experiments.

The wellness score deserves particular caution. A single score can make a complex profile appear more definitive than the underlying evidence permits. Without a documented scoring rubric, validated reference ranges and clinical review, it should be labelled as an AI-generated summary indicator rather than a health grade. Future versions could replace or supplement the score with transparent summaries of the profile fields that informed the output. This would make it easier for parents to understand which reported habits led to a recommendation and reduce the risk of treating a generated number as a diagnosis.

Another limitation concerns the source and completeness of input data. The current prototype depends on information entered by a guardian and does not independently verify the child's measurements or reported behaviour. Missing or inaccurate inputs may influence the response. The report does not describe longitudinal data collection, a comparison group, a pediatric reference dataset or a clinician-reviewed ground-truth set. As a result, the current evidence supports a demonstration of system integration and report generation, but not claims about improved health outcomes, predictive accuracy or superiority over existing clinical services.

A. Evaluation Protocol for a Future Version

A rigorous next-stage evaluation should separate software reliability from the quality of the generated wellness content. Software tests can use synthetic profiles to check required-field validation, accepted input ranges, response parsing, missing JSON keys, API timeouts and recovery from service errors. The results should report the number of test cases, the number that pass each check and the types of failure observed. Performance measurements can include median response time, API error frequency and approximate cost per completed request, with the model version and prompt configuration recorded for reproducibility.

Content evaluation requires a different protocol. Pediatric professionals should review de-identified sample outputs using a predefined rubric covering factual correctness, age appropriateness, relevance to the entered profile, clarity, completeness and potential for harm. Reviewers should be able to mark a recommendation as unsupported or unsuitable and explain why. Parent or guardian usability sessions can examine whether users understand that the wellness score is generated by an AI model, can distinguish general lifestyle suggestions from medical advice, and know when to seek professional care. These evaluations should be conducted with appropriate institutional oversight and informed consent, particularly if any real child information is involved.

The current project report does not contain results from this proposed protocol. It is included here to define a reproducible path from a functional prototype toward a more thoroughly assessed application. Until such evaluation is completed, the system should be described as a proof-of-concept wellness information assistant rather than a clinically validated pediatric decision-support system.

VI. CONCLUSION AND FUTURE WORK

This paper presented HEALTH LLM, a transformer-driven personalized wellness companion for children. The implemented prototype combines a Streamlit interface, Pydantic-based profile validation, structured prompt generation, GPT-4o-mini integration and a categorized recommendation dashboard. It accepts multiple child wellness indicators and produces a consolidated report covering a wellness score, daily plan, nutrition, physical activity, mental-wellness support, risk information and parent/guardian guidance. The project demonstrates an end-to-end software

workflow for presenting AI-generated wellness information in an accessible format.

The present work is limited to functional demonstration using sample profiles and does not establish clinical validity, diagnostic capability or measurable health outcomes. Future work should include evaluation by pediatric healthcare professionals, a predefined test protocol, systematic assessment of factuality and harmful outputs, and usability testing with intended users. Further extensions may include multilingual interaction, voice input, carefully governed wearable-device integration and longitudinal tracking with explicit consent. A safety layer that detects urgent symptom descriptions and directs users to appropriate professional services should also be investigated. Any future deployment involving children should follow applicable privacy, consent and healthcare requirements.

REFERENCES

[1] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention Is All You Need," in Proc. Advances in Neural Information Processing Systems (NeurIPS), vol. 30, 2017, pp. 5998–6008.

[2] T. B. Brown et al., "Language Models are Few-Shot Learners," in Proc. Advances in Neural Information Processing Systems (NeurIPS), vol. 33, 2020, pp. 1877–1901.

[3] R. Miotto, F. Wang, S. Wang, X. Jiang, and J. T. Dudley, "Deep learning for healthcare: Review, opportunities and challenges," Briefings in Bioinformatics, vol. 19, no. 6, pp. 1236–1246, 2018.

[4] K. Singhal, T. Tu, J. Gottweis et al., "Toward expert-level medical question answering with large language models," Nature Medicine, vol. 31, pp. 943–950, 2025, doi: 10.1038/s41591-024-03423-7.

[5] M. Sallam, "The utility of ChatGPT in healthcare education, research, and practice: Systematic review on the promising perspectives and valid concerns," Healthcare, vol. 11, no. 6, Art. no. 887, 2023, doi: 10.3390/healthcare11060887.