International Journal of Engineering Research and Science & Technology

www.ijerst.org

ISSN : 2319-5991

Vol. 22 No. 3 (2026)



ijerst.editor@gmail.com

editor@ijerst.com

ISSN 2319-5991 www.ijerst.com
© 2026 IJERST. All Rights Reserved
Peer Reviewed, Referred & Indexed Journal

Research Paper

HYPOTHESIS GPT

[1] B Divya, [2] K Bhargavi, [3] G Tony Raj, [4] D Akhilesh, [5] K Dileep Kumar
[1]Assistant Professor, [2345]Students
Department of CSE(AIML)
Siddhartha Institute of Technology & Sciences, Narapally
bdivya.cse@siddhartha.co.in, 23TQ1A66B1@siddhartha.co.in, 23TQ1A6682@siddhartha.co.in, 23TQ1A6666@siddhartha.co.in, 23TQ1A66A7@siddhartha.co.in

Abstract

Hypothesis GPT is an Artificial Intelligence-based research assistant designed to help students, researchers, and professionals formulate, refine, and evaluate research hypotheses. Developing a good hypothesis is an important part of scientific and academic research because it provides a clear direction for investigation. However, many researchers face difficulties in converting a broad research idea into a specific, measurable, and testable hypothesis. The proposed system uses AI to simplify this process.

The system accepts a research topic, problem statement, research question, or observation as input. It analyzes the provided information using Natural Language Processing techniques and identifies important concepts, variables, relationships, and research objectives. Based on this analysis, the system generates one or more possible hypotheses in a structured format. It can also identify independent, dependent, and control variables associated with the hypothesis.

Hypothesis GPT can generate different types of hypotheses, including directional, non-directional, null, alternative, correlational, and comparative hypotheses. The system explains each generated hypothesis in simple language and identifies the variables involved. It can also provide suggestions for making a hypothesis more specific, measurable, and suitable for empirical testing.

The system can further assist researchers by evaluating the logical structure of a proposed hypothesis. It can identify whether the hypothesis is clearly stated, testable, measurable, relevant to the research question, and supported by the provided context. The system can suggest improvements when the hypothesis is too broad, ambiguous, or difficult to test. It can also recommend possible research methods and measurable indicators without claiming that the hypothesis has been scientifically proven.

Overall, Hypothesis GPT aims to make research planning faster, clearer, and more accessible. It can be useful for students preparing academic projects, researchers developing research proposals, and professionals exploring new ideas. Future enhancements can include integration with academic databases, citation-based evidence retrieval, statistical test recommendations, experiment-design assistance, and automated hypothesis evaluation using research literature.

I. INTRODUCTION

Research begins with identifying a problem, asking meaningful questions, and developing ideas that can be investigated systematically. A hypothesis provides a proposed explanation or prediction that can be examined using observations, experiments, surveys, or statistical analysis. Developing a good hypothesis requires clear understanding of the research problem and the variables involved. Students and new researchers may find this process challenging.

Traditional hypothesis development usually involves reading research papers, identifying gaps, understanding existing theories, and manually formulating possible relationships between variables. Researchers may spend considerable time refining the wording of their hypotheses. A poorly structured hypothesis can make research objectives unclear and may create difficulties during data collection and analysis. Therefore, an intelligent tool that assists researchers during the early stages of hypothesis development can be useful.

Artificial Intelligence and Natural Language Processing provide new opportunities for supporting research activities. AI systems can analyze text, extract concepts, identify relationships, summarize information, and generate structured content. Generative AI can use these capabilities to transform a research topic or question into possible hypotheses. However, AI-generated hypotheses should be treated as suggestions that require human review rather than automatically accepted scientific conclusions.

Hypothesis GPT is designed to act as an AI-powered research assistant. A user can enter a research topic, problem statement, research question, or initial idea. The system analyzes the input and identifies key concepts, variables, relationships, and objectives. It then generates possible hypotheses and explains their structure. Users can review the suggestions and modify them according to their research requirements.

The proposed system focuses on assisting researchers rather than replacing human reasoning. It provides structured suggestions, identifies potential variables, checks basic testability, and helps improve the clarity of hypotheses. By combining AI-generated suggestions with human validation, Hypothesis GPT can support better research planning while allowing the researcher to make the final decision about the hypothesis.

II. LITERATURE SURVEY

Traditional research methodology emphasizes hypothesis development as an important stage of scientific investigation. Researchers generally begin with a research problem or question and examine existing theories and literature before developing a hypothesis. This process requires subject knowledge and careful reasoning. While traditional methods can produce high-quality hypotheses, they may require significant time and extensive literature review.

Research support systems have been developed to assist researchers with literature discovery, information retrieval, citation management, and knowledge organization. These systems help users identify relevant research papers and understand existing work. However, most traditional research tools focus on retrieving or organizing information rather than automatically transforming research questions into structured hypotheses.

Natural Language Processing has been applied to academic text analysis, including keyword extraction, topic modeling, entity recognition, relation extraction, and document classification. These techniques can identify important concepts and relationships from research papers. NLP-based systems can therefore provide useful inputs for hypothesis-generation applications. However, conventional NLP techniques may have difficulty understanding complex relationships and research context.

Machine learning techniques have also been explored for discovering patterns and relationships in research data. Data-driven approaches can identify correlations and potentially useful variables. However, identifying a statistical relationship does not automatically establish a valid scientific hypothesis. Human researchers still need to interpret findings, establish theoretical reasoning, and determine whether a proposed relationship is meaningful and testable.

Recent advances in **Generative AI and Large Language Models** have created new possibilities for research assistance. These models can generate research questions, summarize papers, identify potential variables, and formulate possible hypotheses using natural language. They can also explain concepts and suggest alternative formulations. However, Generative AI can produce unsupported or inaccurate statements, so human review, reliable sources, and validation mechanisms are important.

Existing research indicates that AI can support researchers during ideation and hypothesis development, but challenges remain in maintaining scientific validity, avoiding unsupported assumptions, identifying measurable variables, and distinguishing suggestions from established findings. The proposed **Hypothesis GPT** addresses these challenges by combining NLP-based research analysis, AI hypothesis generation, variable identification, testability checking, explanation, and human review into a single research-support system.

III. SYSTEM ANALYSIS

Hypothesis GPT is designed to assist users in generating and refining research hypotheses from a research topic or question. The user first enters a research problem, topic, observation, or research question through the system interface. The input-processing module cleans and organizes the submitted information. NLP techniques identify important concepts, entities, variables, and relationships within the research context. The system determines possible independent, dependent, and control variables from the identified concepts. An AI reasoning module analyzes the relationship between the variables and generates candidate hypotheses. Multiple hypothesis types such as null, alternative, directional, and non-directional hypotheses can be produced. A validation module checks whether each hypothesis is clear, relevant, measurable, and potentially testable. The system explains the reasoning and structure behind each suggested hypothesis. Users can edit, compare, refine, or regenerate hypotheses according to their research requirements. The system can also suggest measurable indicators and possible research approaches. Finally, the researcher selects and validates the most appropriate hypothesis for further investigation.

Existing System

Traditional hypothesis development is mainly performed manually by researchers, students, and academic supervisors. The researcher first identifies a research problem and conducts a literature review to understand existing knowledge. Important variables and relationships are then identified from previous studies and theoretical frameworks. The researcher manually develops one or more possible hypotheses based on this information. Refining the wording of a hypothesis may require multiple revisions and discussions with supervisors or experts. Beginners may find it difficult to determine whether a hypothesis is sufficiently specific and testable. Traditional research tools can help find papers but generally do not automatically convert research questions into hypotheses. Researchers may also need to manually identify independent and dependent variables. Comparing multiple hypothesis alternatives can require additional effort. There is limited automated feedback on whether a hypothesis is measurable or logically connected to the research question. As a result, the traditional process can be time-consuming, particularly for students and researchers who are new to research methodology.

Disadvantages of Existing System

- Requires significant manual effort.
- Hypothesis development can be time-consuming.
- Beginners may struggle to formulate testable hypotheses.
- Limited automated identification of research variables.
- Requires extensive literature review before formulation.
- Limited instant feedback on hypothesis quality.
- Difficult to generate multiple alternatives quickly.
- Manual refinement may require repeated revisions.
- Limited automated checking of testability.
- Research tools generally focus on literature rather than hypothesis generation.
- Difficult to identify relationships between variables automatically.
- Depends heavily on researcher experience and subject knowledge.

Proposed System

The proposed **Hypothesis GPT** uses Artificial Intelligence and Natural Language Processing to assist researchers in generating clear and testable research hypotheses. The user provides a research topic, question, problem statement, or initial observation as input. The system analyzes the input and extracts important concepts, variables, objectives, and relationships. An AI research-analysis module identifies potential independent, dependent, and control variables. The hypothesis-generation engine creates multiple candidate hypotheses based on the identified relationships. The system can generate null, alternative, directional, non-directional, correlational, and comparative hypotheses. Each hypothesis is presented with an explanation of the variables and expected relationship. A validation module evaluates clarity, relevance, measurability, specificity, and basic testability. The system identifies vague or unsupported assumptions and provides suggestions for improvement. Users can compare different hypotheses, edit their wording, and regenerate alternatives. The system can also suggest measurable indicators and possible research approaches for further consideration. The final hypothesis is selected and validated by the researcher before being used in actual research.

Advantages of Proposed System

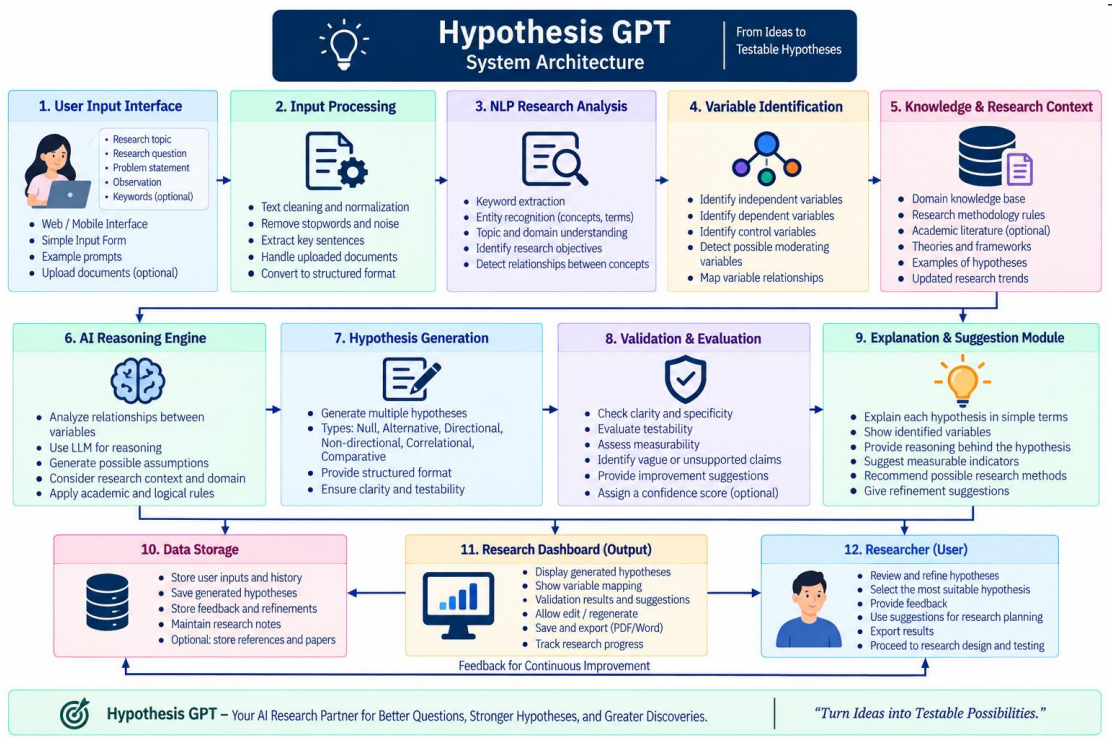
- Generates research hypotheses quickly.
- Reduces the effort required during initial research planning.
- Identifies potential research variables automatically.
- Generates multiple hypothesis alternatives.
- Supports different types of hypotheses.
- Provides explanations for generated hypotheses.
- Checks basic clarity and testability.
- Helps beginners understand hypothesis structure.
- Allows users to refine and regenerate hypotheses.
- Suggests measurable variables and indicators.
- Improves research planning efficiency.
- Keeps the researcher involved in final validation.

IV. METHODOLOGY

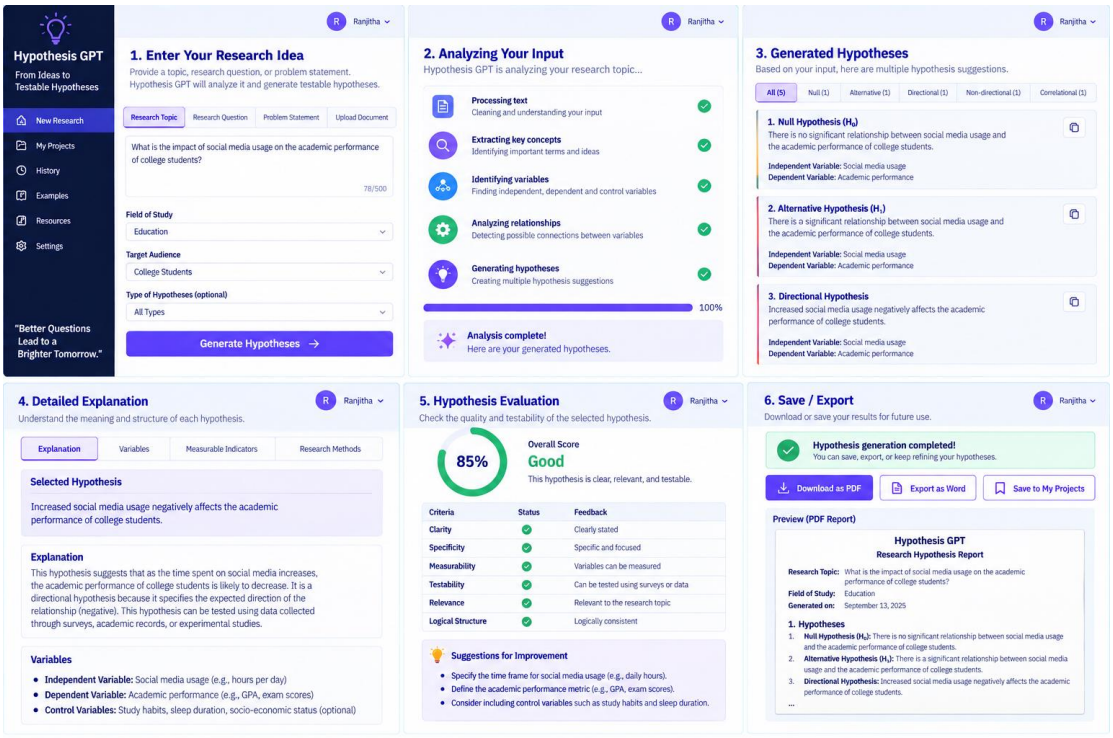
The methodology begins when the researcher enters a topic, research question, problem statement, or observation into the system. The input is first processed using Natural Language Processing techniques to remove unnecessary information and identify important concepts. The system extracts keywords, entities, research objectives, and potential variables from the input. A semantic analysis module examines relationships between the identified concepts. The variable identification module classifies possible independent, dependent, and control variables. The AI reasoning engine analyzes these variables and generates possible relationships that could be investigated. A Generative AI module converts these relationships into clear and structured hypotheses. The system generates different hypothesis types depending on the research requirement. A validation module evaluates each hypothesis for clarity, specificity, relevance, measurability, and basic testability. The system provides explanations and suggestions for improving weak or ambiguous hypotheses. The researcher reviews, edits, compares, and selects the most suitable hypothesis. Finally, the approved hypothesis can be used as a starting point for research design, data collection, and statistical testing.

SYSTEM ARCHITECTURE

The Hypothesis GPT architecture consists of several interconnected modules that support the complete hypothesis-development process. The User Interface allows researchers to enter research questions, topics, or problem statements. The Input Processing Module cleans and normalizes the submitted research information. The NLP Research Analysis Module identifies keywords, concepts, entities, objectives, and relationships. The Variable Identification Module detects potential independent, dependent, and control variables. The Knowledge and Research Context Module stores relevant research concepts, methodological information, and optional literature references. The AI Reasoning Engine analyzes relationships between variables and generates possible research assumptions. The Hypothesis Generation Module produces different types of candidate hypotheses. The Validation Module evaluates hypotheses for clarity, specificity, relevance, measurability, and testability. The Explanation Module describes the variables and relationships behind each hypothesis. The Refinement Module allows users to edit, compare, and regenerate hypotheses.



V. RESULT AND OUTPUT



VI. CONCLUSION

The Hypothesis GPT system provides an intelligent and efficient solution for assisting students and researchers in developing clear, relevant, and testable research hypotheses. By accepting a research topic, question, problem statement, or observation, the system can analyze the input and generate suitable hypothesis suggestions.

The system combines Artificial Intelligence, Natural Language Processing, semantic analysis, and Generative AI to identify important concepts, variables, and relationships. It can generate different types of hypotheses, including null, alternative, directional, non-directional, correlational, and comparative hypotheses, along with explanations of the variables involved.

The validation component helps evaluate generated hypotheses based on clarity, specificity, relevance, measurability, and basic testability. Researchers can review, edit, compare, and refine the suggestions before selecting a final hypothesis. This ensures that AI supports the research process while the researcher remains responsible for the final scientific decision.

Overall, Hypothesis GPT can reduce the time and effort required during the early stages of research planning and make hypothesis development easier for beginners. It can be useful for students, academic researchers, and professionals. Future enhancements can include integration with academic research databases, citation-based evidence analysis, statistical test recommendations, experiment-design assistance, and stronger literature-based hypothesis validation.

References

[1] Kumar, R. D., Prudhviraj, G., Vijay, K., Kumar, P. S., & Plugmann, P. (2024). Exploring COVID-19 through intensive investigation with supervised machine learning algorithm. In Handbook of Artificial Intelligence and Wearables (pp. 145-158). CRC Press.
[2] Swathi, B., Vijay, K., Sushanth Babu, M., & Dinesh Kumar, R. (2024, November). Machine Learning Techniques in Cloud Based Intrusion Detection. In The International Conference on Artificial Intelligence and Smart Environment (pp. 557-564). Cham: Springer Nature Switzerland.
[3] Sv satyakrishna, shirisha rangu ,bhargavi nalacheruve.(2024) Prospective investigation on colorectal cancer with SMOTE on machine learning Algorithm
[4] Dr.G.Vishnu Murthy, BhargaviNalacheruve 1Professor, Department of computer Science & engineering, Anurag University, TS, India. 2Student, Department of computer Science & engineering, Anurag University, TS, India.
[5] V. N. S. Manaswini, K. K, C. Nigam, S. S. Ali, R. Niranjana, and Suman, "Real-Time Object Detection in Drone Surveillance Using YOLOv5," in Proc. 2025 3rd Int. Conf. IoT, Communication and Automation Technology (ICICAT), Gorakhpur, India, 2025, pp. 1–6, doi: 10.1109/ICICAT68430.2025.11414670.
[6] B. Soundarya, V. N. S. Manaswini, M. Ayyakrishnan, R. D. Kumar, "Contextual Analysis of Big Data Analytics in Intelligent Transportation Frameworks," in Intersection of Artificial Intelligence, Data Science, and Cutting-Edge Technologies: From Concepts to Applications in Smart Environment, Lecture Notes in Networks and Systems, vol. 1353, Cham: Springer, 2025, doi: 10.1007/978-3-031-88304-0_79.
[7] R. D. Kumar, V. N. S.Manaswini, "Applications of blockchain in smart cities: detecting fake documents from land records using blockchain technology," in Blockchain for Smart Cities, Elsevier, 2021, pp. 105–117, doi: 10.1016/B978-0-12-824446-3.00017-X.
[8] Tejavath Veeramma, Badarla Anil, Guguloth Ravinder, "An advanced movie recommender using collaborative filtering and sentiment analysis," International

Research Journal of Modernization in Engineering Technology and Science, vol. 7, no. 7, July 2025, doi: 10.56726/IRJMETS81618.

[9] Ravi Kumar Banoth, Ramana Murthy B V, "Automatic crop recommendation system using LightGBM and decision tree machine learning models," Journal of Machine and Computing, vol. 5, no. 1, pp. 343, Jan. 2025, doi: 10.53759/7669/jmc202505026.

[10] Ravi Kumar Banoth, Dr. B.V. Ramana Murthy, "Smart agriculture through IoT and machine learning for analyzing carbon footprints," in Proc. Int. Conf. Computer Science and Communication Engineering (ICCSCE), Apr. 2025.

[11] Ravi Kumar Banoth, B. V. Ramana Murthy, "Soil image classification using transfer learning approach: MobileNetV2 with CNN," SN Computer Science, vol. 5, art. no. 199, 2024, doi: 10.1007/s42979-023-02500-x.