

Assessing ChatGPT for Aeronautical Engineering Applications

Background – The rise of Natural Language Processing Models, especially ChatGPT has caused heated debate in academic circles, due to being able to produce answers that in many domains surpass those that students can give, while sometimes missing the mark by a wide margin.

Purpose – This project assesses the reliability, accuracy, and truthfulness of such a tool regarding aeronautical engineering questions. To this end, a classification and evaluation scheme was devised, to numerically assess the viability of using currently available models for questions about several disciplines of aircraft design and engineering.

Method – After a general evaluation of ChatGPT, questions will be assigned one of three tiers: TIER I: Explanatory. Questions whose answer is an explanation, rather than a design decision or a calculation. The answers will be graded based on their legibility, accuracy and whether a disambiguation has taken place. TIER II: Evaluation. Questions whose answer is a design decision. These answers will be graded based on their exhaustiveness, accuracy and again, disambiguation. TIER III: Calculation. Questions that are solved by calculation. Given certain input values, the model will be graded on the accuracy of its output as compared to a reference value. While ChatGPT does not itself do calculations beyond a certain complexity, it does give formulas which then can be used to arrive at a value.

Findings – *General findings:* Accuracy: In general, ChatGPT 3.5 demonstrated high accuracy in providing correct and relevant information for aeronautical engineering queries. Relevance: The responses were mostly relevant to the queries, with occasional off-topic or ambiguous answers. Speed: ChatGPT 3.5 provided quick responses, making it efficient for time-sensitive tasks. Usability: The model's interface was user-friendly, and its responses were easy to understand, even for complex engineering queries. Design Assistance: ChatGPT 3.5 can assist engineers in brainstorming design ideas, providing information on materials, aerodynamics, and structural analysis. Data Analysis: The model can aid in data interpretation and analysis, helping engineers make informed decisions. Education and Training: ChatGPT 3.5 can serve as a learning tool for students and professionals, offering explanations and insights into aeronautical concepts and principles. Specialized Knowledge: While knowledgeable, ChatGPT 3.5 lacks specialized expertise in certain niche areas of aeronautical engineering. Ambiguity: Occasionally, the model provided ambiguous or incomplete answers, requiring further clarification. Dependence on Data Quality: The accuracy and relevance of responses heavily depend on the quality of the data provided to the model. *Findings related to the TIERS:* ChatGPT 3.5 is performing quite well on problems defined as TIER I, while doing less well on TIER II. Problems defined as TIER III were not solve to any degree of satisfaction.

Value – The project gives potential users in aeronautical engineering a realistic view about what can be expected from ChatGPT 3.5 depending on the type of problem and the specialized area within aeronautical engineering.

This informative poster is based on a student project with the same title.

This is an abstract answering the Call for Papers of the German Aerospace Conference 2024 for an informative poster at the conference.

Prof. Dr.-Ing. Dieter Scholz, MSME
Hamburg University of Applied Sciences
Department of Automotive and Aeronautical Engineering
Aircraft Design and Systems Group (AERO)
<http://www.ProfScholz.de>
info@ProfScholz.de