Evaluating Engineering Design Descriptions using AI-based Chatbot

Abstract

This research introduces an AI-driven chatbot designed to evaluate engineering design descriptions based on novelty, feasibility, and validity, addressing the challenges of efficient and accurate assessment. Leveraging fine-tuned BERT base uncased models, the chatbot assesses user inputs through a three-step process, achieving accuracies of 80-85%. The system's architecture and integration of AI models streamline evaluation, offering a rapid alternative to traditional human-based methods. The study highlights the significance of tailored data for model training and emphasizes the potential of AI-based evaluation systems in diverse domains.

Introduction

The advent of artificial intelligence has revolutionized numerous aspects of our lives, transcending various industries and domains. Among these, the application of AI in natural language processing (NLP) stands out as a transformative force. In the realm of engineering design, the evaluation of design descriptions is a crucial step that often requires a significant investment of time and resources. This research delves into the development of an AI-based chatbot tailored to this specific challenge.

The motivation behind this research stems from the need for efficient and accurate evaluation of engineering design descriptions. Traditionally, this task has been performed through labor-intensive manual assessments, which are not only time-consuming but also prone to subjectivity. The emergence of powerful language models, such as BERT base uncased, presents an opportunity to automate and enhance this process through objective and data-driven means.

The primary objective of this research is to create an AI-based chatbot that can evaluate design descriptions based on three crucial dimensions: novelty, feasibility, and validity. By training BERT models on datasets curated from engineering design surveys, this chatbot aims to replicate and, in certain aspects, surpass human evaluative capabilities. The significance of this endeavor lies not only in its potential to enhance efficiency but also in the novel approach to evaluation itself.

This paper proceeds with an exploration of the research's background, methodologies, results, and insights. It delves into the intricacies of data collection, model fine-tuning, and system implementation. Through a detailed analysis of the achieved accuracies and the lessons learned, this paper sheds light on the significance of data quality, model selection, and architectural innovation in the development of AI-driven evaluative systems.

Literature Review:

- ARE YOU FEELING HAPPY? THE EFFECT OF EMOTIONS ON PEOPLE'S INTERACTION EXPERIENCE TOWARD EMPATHETIC CHATBOTS
- A Meta-Analysis of Factors Affecting Trust in Human-Robot Interaction
- Machines and Mindlessness: Social Responses to Computers
- Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: A mixed-methods study
- Machine learning-based design concept evaluation
- Product design concept evaluation using rough sets and VIKOR method
- An integrated AHP and VIKOR for design concept evaluation based on rough number
- Evaluation of product design concepts using grey-fuzzy information axiom

Methodology:

In this section, we outline the precise methodology employed to develop and implement the AI-powered chatbot for engineering design evaluation.

Chatbot Architecture:
The architecture of the AI-powered chatbot was meticulously designed to ensure a seamless user experience. When a user initiates interaction, the chatbot commences with a cordial greeting. Subsequently, the user is prompted to provide three fundamental inputs: the design's title, description, and objective. These inputs collectively serve as pivotal components in the subsequent assessment process.

Utilizing BERT Models:
At the heart of the chatbot's evaluation prowess lies the utilization of BERT (Bidirectional Encoder Representations from Transformers) models. These models, based on transformer architecture, possess the capability to grasp contextual nuances within text data. To tailor these models to our specific task, we embarked on a process of fine-tuning. We accomplished this by employing specialized datasets derived from comprehensive surveys encompassing engineering designers' descriptions of washing machine designs.

Three-Step Assessment Protocol:
The core of the chatbot's functionality is its three-step assessment protocol. Upon receiving the user's design inputs, the chatbot engages in a sequential assessment process involving novelty, feasibility, and validity evaluation. The description provided by the user undergoes scrutiny through these distinct BERT models, each aligned with its respective assessment criterion. The outcomes of these assessments cumulatively guide the chatbot's decision-making process regarding the design's suitability.

User Interaction Dynamics:

In instances where the design description fails to meet the established criteria, the chatbot strategically employs fallback messages to prompt users for revised submissions. These iterative interactions facilitate a dynamic and collaborative engagement between the user and the chatbot. Furthermore, motivational responses are integrated to encourage users in refining their design descriptions, thereby enhancing the potential for successful evaluation.

Title: Evaluating Engineering Design Descriptions using AI: A Novel Approach to Assessing Novelty, Feasibility, and Validity

Abstract:

- Brief overview of the research project, its objectives, methodologies, and key findings.
- Emphasis on the significance of the AI-powered chatbot for evaluating engineering design descriptions based on novelty, feasibility, and validity.
  Introduction
    - Introduction to the research area: AI in engineering design evaluation.
    - Statement of the problem: Challenges in efficient and accurate assessment of engineering design descriptions.
    - Purpose of the study: Developing an AI chatbot for evaluating novelty, feasibility, and validity.
    - Research questions and objectives: How AI models assist and enhance in evaluating engineering design descriptions.

Literature Review:

- Overview of AI applications in engineering and design.
- Previous research on using AI models for text analysis and assessment.
- Existing methods for evaluating novelty, feasibility, and validity of design descriptions.
- Comparison of AI-based evaluation with traditional human-based evaluation approaches.

Methodology:

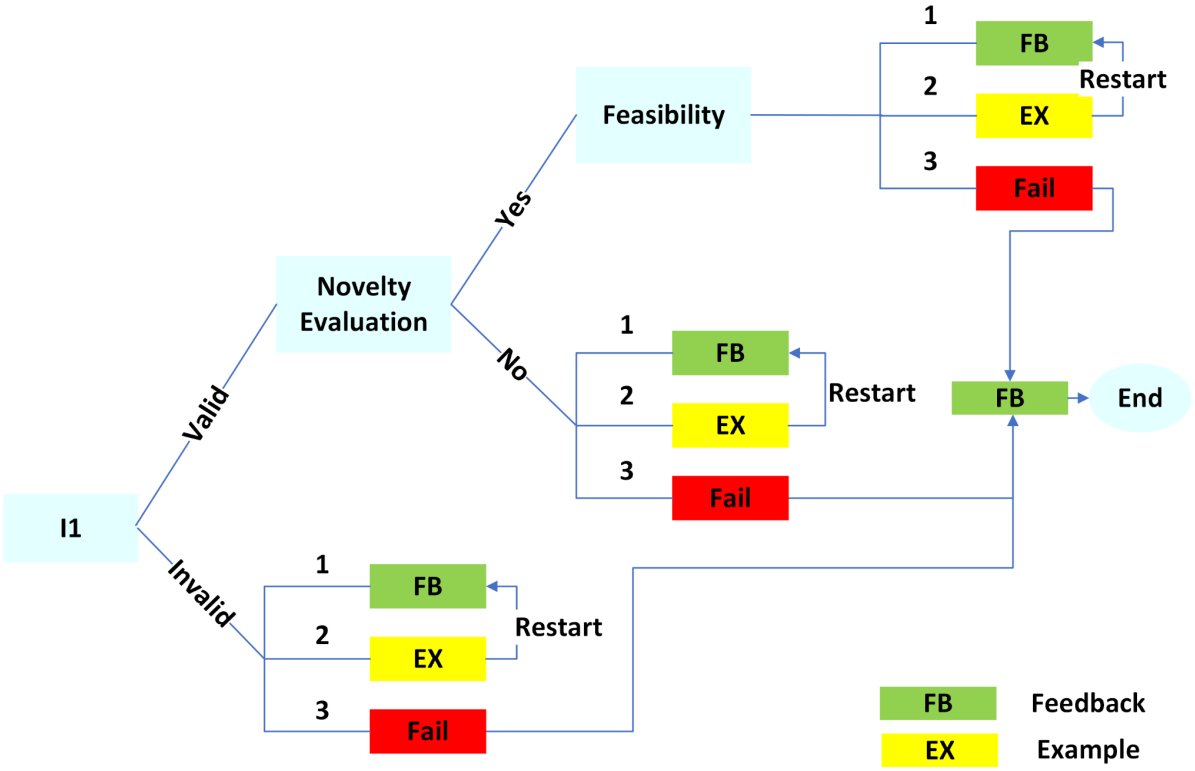
- Description of the AI chatbot architecture:
    - Three-step assessment process: Novelty, Feasibility, Validity.

- Integration of BERT base uncased models for each assessment.
- User input requirements: title, description, objective.

- Data collection and preprocessing:
  - Creation of datasets using engineering design descriptions from surveys.
  - Dataset descriptions - Novelty, Feasibility and Validation
  - Using few-shot prompting on ChatGPT to generate further data
  - Preprocessing techniques to prepare data for model training.
- Model training and fine-tuning:
  - Details of fine-tuning BERT base uncased models on respective datasets.
  - Evaluation metrics used to achieve 80-85% accuracies for each model.

System Implementation

- Technical details of implementing the AI chatbot system.
  - Utilising Flask Back-End, with Hugging Face Models and React for the Front-End
- User interaction flow:
  - Greeting and user input requirements.
  - Passing description through novelty, feasibility, and validity models.
  - Decision-making process based on assessment outcomes.
- Handling user failures:
  - Fallback messages and motivational responses.
  - System termination in case of repeated failures.



Results and Discussion

- Presentation of accuracy results for each model: novelty, feasibility, validity.
- Analysis of system performance across different types of design descriptions. (Short, long, correct, incorrect)
- Comparison of AI model evaluation with human evaluation in terms of time and accuracy **(We need to figure this out but I believe this is crucial)**
- Discussion on limitations and challenges faced during system implementation (Lack of data)

Implications and Future Work -

- This system can be used to create AI powered evaluation applications over various domains
- Implications of the AI-powered chatbot for engineering design evaluation.
- Potential applications in other domains and industries.
- Future enhancements and refinements to the chatbot system:
    - Incorporation of more advanced NLP models.
    - Expansion of assessment criteria beyond novelty, feasibility, and validity.

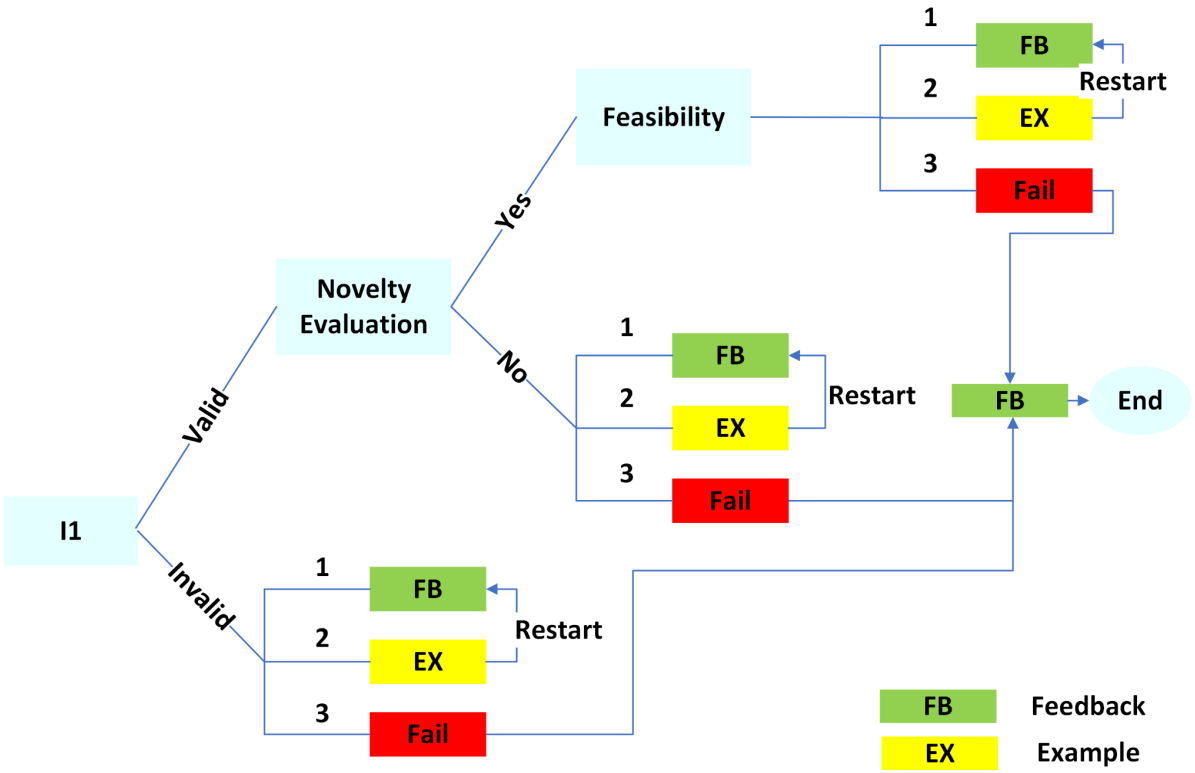
Conclusion

- Recap of research objectives and key findings.
- Contributions of the study to the field of engineering design evaluation.
- Closing remarks on the potential of AI-driven assessment systems.

References:

1. Liao, Ting, and Bei Yan. "Are You Feeling Happy? the Effect of Emotions on People's Interaction Experience Toward Empathetic Chatbots." *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. Vol. 86236. American Society of Mechanical Engineers, 2022.
2. Hancock, Peter A., et al. "A meta-analysis of factors affecting trust in human-robot interaction." *Human factors* 53.5 (2011): 517-527.
3. Nass, Clifford, and Youngme Moon. "Machines and mindlessness: Social responses to computers." *Journal of social issues* 56.1 (2000): 81-103.
4. Nadarzynski, Tom, et al. "Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: A mixed-methods study." *Digital health* 5 (2019): 2055207619871808.
5. Camburn, Bradley, et al. "Machine learning-based design concept evaluation." *Journal of Mechanical Design* 142.3 (2020): 031113.
6. Tiwari, Varun, Prashant Kumar Jain, and Puneet Tandon. "Product design concept evaluation using rough sets and VIKOR method." *Advanced Engineering Informatics* 30.1 (2016): 16-25.
7. Zhu, Guo-Niu, et al. "An integrated AHP and VIKOR for design concept evaluation based on rough number." *Advanced Engineering Informatics* 29.3 (2015): 408-418.

8. Akay, Diyar, and Osman Kulak. "Evaluation of product design concepts using grey-fuzzy information axiom." *2007 IEEE International Conference on Grey Systems and Intelligent Services*. IEEE, 2007.



Chat Flow

Naveen Mathews Renji - | September 26, 2023