Proposal for GUS Academic Summit

On 19th and 20th May 2025

Title: Effectiveness of ChatGPT in Coding: A Comparative Analysis of Popular LLMs

Speaker: Prof. Dr. Rand Kouatly

Abstract:

The study explores the effectiveness and efficiency of the new Large Language models, as OpenAI model ChatGPT powered by GPT-3.5 and GPT-4, Google Gemini, etc, in programming tasks to understand its impact on programming and, potentially, software development. To measure the performance of these models, a quantitative approach was employed using the Mostly Basic Python Problems (MBPP) dataset. In addition to the direct assessment of GPT-3.5 and GPT-4, a comparative analysis involving other popular large language models in the AI landscape, notably Google's Bard and Anthropic's Claude, was conducted to measure and compare their proficiency in the same tasks. The study identifies the challenges and limitations of using large language models in programming. The breakthrough in automated code generation has been significantly propelled.

The results highlight the strengths of ChatGPT models in programming tasks, offering valuable insights for the AI community, specifically for developers and researchers. As the popularity of artificial intelligence increases, this study serves as an early look into the field of AI-assisted programming.