

CS 490 Proposals

1 Analyzing the Effectiveness of GitHub CoPilot (2 students)

GitHub Copilot (Your AI programmer) has recently been released (<https://copilot.github.com>) GitHub Copilot is an AI pair programmer that helps you write code faster and with less work. GitHub Copilot draws context from comments and code, and suggests individual lines and whole functions instantly. GitHub Copilot is powered by OpenAI Codex, a new AI system created by OpenAI.

We would like to explore the limits of GitHub CoPilot by conducting several experiments.

2 Data science for Software Engineering (1 student)

Software analytics aims to obtain insightful and actionable information from software artifacts that help practitioners accomplish tasks related to software development, systems, and users.” [1] With software analytics, the main idea is mining the big data in multiple software repositories to produce insights for decision makers in the software development process.

The project has the following stages;

Getting familiar with the topic: This requires some guided reading about Software Analytics and Data science.

Research: First step is finding the proper software analytics question (s) that have practical impact to software practitioners. Then, you will investigate /combine various data mining techniques (machine learning algorithms, deep learning, big data graph visualization etc.) to answer the software analytics question.

Implementation: Application of the right data science technique to the right data science question

Experimentation: You will evaluate your approach based on available open source repositories.

[1] D. Zhang, S. han, Y. Dan, J.-G. Lou, H Zhang: ”Software Analytics in Practice”. IEEE Software, Sept./Oct. 2013, pp. 30-35.

Please check <http://www.cs.bilkent.edu.tr/~korpe/courses/CS490/index.html> for eligibility. If you are interested, please send me an email with your CV and transcript (eraytuzun@cs.bilkent.edu.tr)