1. Goal-oriented Automated Code Generation (Taken)

In recent years, there has been significant interest in program synthesis, in which the goal is to generate programs from high-level specifications automatically. Automated program synthesis has proven to be useful to both end-users and programmers: for instance, programming-by-example (PBE) has been used to automate tedious tasks that arise in everyday life, such as string and format manipulations in spreadsheets or data wrangling tasks on tabular and hierarchical data. Program synthesis has also been used for improving programmer productivity by automatically completing parts of a program or helping programmers use complex APIs.

This project aims to automatically generate source code that could achieve pre-defined goals provided by developers. We plan to start by searching for goal related code snippets from thousands of open source projects hosted on GitHub. We will then apply reinforcement learning techniques to learn steps that could transfer candidate code snippets into a program that could achieve the defined goal.

2. Empirical Study on Cross-Platform Mobile Development Frameworks (Taken)

Cross-platform app development frameworks allow mobile app developers to develop the same app on different operating systems, i.e., android and iOS. It is extremely attractive for developers as it will save their time and human efforts in maintaining different code with the same function on two platforms.

Given the multiple frameworks including Flutter, Xamarin, React Native etc. available in market, this project aims to conduct an empirical study to understand how developers discuss about cross-platform mobile development frameworks on question answering sites like Stack Overflow and forums. We will propose a model to filter out discussions related to multiple cross-platform mobile development frameworks and provide insight for various stakeholders.

3. Mining Data Analysis Competition Code

Kaggle is the most well known and the largest competition platform for predictive data modeling and analytics.

With thousands of ongoing or completed data analytics competitions on Kaggle, this project aims to analyze patterns existing in the notebook tutorials provided by competition participants and understand what makes a good data analytics code in

terms of performance and quality. We are also interested to mine common pipelines for data analytics to guide people who are interested in learning data analytics. It is also interesting to investigate if some data analytics code can be automatically generated.

4. Course Recommender System for Queen's School of Computing Students
Queen's school of computing has provided tones of plans for students to customize their course schedule during four years' study. However, having too many choices sometimes exhaust us. It is a time-consuming process for students to select a series of courses based on their own interest, course prerequisite constraints, and degree requirements. Besides, with the overwhelming number of students requiring supervision in course selection, related SOC faculty members and admin staffs need to spend increasing time on consulting service for course selection.

 This project aims to propose a novel constraint-based course recommendation system for SOC based on our degree plans and course prerequisite relationship. We hope such a system could help SOC students better select courses and reduce SOC staff's consultation time on repeated and trivial course recommendation tasks,