**Project done by:**

**Cyprian Nwachukwu**

**Apekshya Rimal**

## **Project Proposal: Replication and extension of Empirical Study on the Relationship Between Developer’s**

**Working Habits and Efficiency paper**

**Paper Description:**

The paper is an empirical study on the impact of software developer work habits (based on the total work time throughout the week, continuous working time and particular period of the day) on efficiency of the work done. The paper addressed three research questions: “Does the day of the week affect development time and work efficiency”? This question attempts to find correlation between software developer efficiency and day of the week, “Is there any relationship between development efficiency and 24-hour work period”? This research question attempts to identify times of the day associated with higher and lower development efficiency, and “does continuous working time impact development efficiency”, this attempts to establish relationship between continuous work time and decline in development efficiency. The authors used interactive dataset from the KaVE project FeedBaG++ plug-in in Visual Studio IDE. The plug-in records data about development activities on the IDE. The authors used chi-square test and residual analysis to test the difference between in ratios for days of week, time ranges and continuous work time ranges for each of the research questions. The research result may be useful to team leads and software development managers to effectively plan their work time for efficient software product delivery.

**Planned Extension**:

We have planned to extend the paper in two ways.

1. We want to use the same categories of time of the day, days of the week and number of successful builds but analyze them differently using regression models to see how software developer productivity is affected.
2. Secondly, we want to include more categories of the KAVE dataset like the “Active Window” and “Active Document” to see if the developer is actually spending time coding and try to find out the reason behind the software productivity.

We will like to extend the paper by building regression models to analyze the dataset differently.

**Research Questions**:

How is software developer productivity affected by different work period of time of day, day of week and continuous work time ranges?

Is the time that the developer actually spends on coding is directly proportional to the developer’s productivity?

**Data source**:

The dataset for the research was interactive dataset from KaVE project’s FeedBaG++ tool used in the MSR 2018 mining challenge. The FeedBaG++ is a plug-in in Visual Studio Integrated Development Environment (IDE). It includes ActivityEvent which records developer actions on the IDE. There are also Build and TestRun events that provide data related to build and run of a program on the IDE. The dataset contains events submitted by FeedBaG++ users. The dataset contains data shared by 81 different developers of different skill levels and backgrounds and the data represents a total of 15,000 hours of work by developers. The authors sampled 328,667 separate events from 11,122,103 events dataset.

Data Source: http://www.kave.cc/datasets

**Reference**:

Rodriguez, Ariel, Fumiya Tanaka, and Yasutaka Kamei. "Empirical study on the relationship between developer's working habits and efficiency." *Proceedings of the 15th International Conference on Mining Software Repositories*. ACM, 2018.