

## **Project Progress Report:**

**Student Name: Maria Fernandez**

My project is to reproduce the paper: Generating Semantic Annotations for Frequent Patterns with Context Analysis

The progress I have made:

1. I have read through the paper a few times, making sure I understand the concepts that are used in the paper, like Pattern, Transactions, Context Unit, Closed Frequent Pattern, Semantically Similar Patterns, etc. I tried focusing in understanding the section that explains the One-pass Microclustering algorithm since this is the part that I will have to work from scratch.
2. I found the dataset that I will be using since the original location is not available in the location anymore, but this is the same data set, just a newer version of it. I made sure it has the data we need.
3. I looked into python libraries to parse XM.
4. I looked into python libraries for Frequent Patterns and for Close Patterns. It is allowed to use the libraries to pre-process the data set, and create the input to the microclustering algorithm. I still have to decide what I am using here.

Remaining Tasks: I have to decide what libraries I am using for Frequent Pattern and Close Pattern and I need to start coding.

Any Challenges: I am working alone because with work and family responsibilities it would be hard to collaborate with others, so it is challenging to take decisions by myself. So far I am not sure they libraries I found are the best ones to use, so I am planning to ask the question in Piazza and see if the other people doing the same paper has found the same libraries or if they have other suggestions.