CSCI 6441
Project Proposal Review

Reviewer: Ryan Gross
Proposer: Musadig Aliyev

Overall this is a very interesting research topic. It is both compelling and practical. When you start doing research on bundling techniques I would recommend trying to also find computer science research papers of others attempting to use machine learning for bundling. You mentioned that today it is very manual, so the existing computer science literature may be sparse, but even if you could find someone attempting to use machine learning to bundle in other industries, like cable packages, that could be a good citation.

Regarding risks, I agree that it will be hard to find real data. However, I don't view this as an issue. Being able to generate data to simulate specific customer profiles that you want to target will be just as convincing. The real risk for this project is being able to quantify how much better the machine learning approach is over the manual approach. Increasing turnover and revenue were cited as performance metrics, but it's not clear to me that you will be able to measure this in your research. It's also not clear that you will have an existing state-of-the-art metric to beat. Some research is hard to measure performance, and I am sure this project will be convincing regardless. However, if you can find some way to establish state-of-the-art performance metrics, and measure the performance of your algorithms numerically to see if they beat the state-of-the-art, that would make the final results even more impressive. People tend to like "black and white" solutions, so if you can numerically claim you are better, no one can refute that. More importantly your customer will have a boss, and your customer will need to convince their boss that it is worth spending money to incorporate machine learning techniques into their business. Having a numerically superior claim only helps.

For a mid-term milestone you mention demonstrating bundling techniques, and then the final milestone is showing these bundling techniques using machine learning. I realize the bundle techniques are new to you, so you have to spend time understanding how they work. However, if a customer were paying you to do this research project they would want to see some machine learning results by the mid-way point. Imagine the customer is a bank. They already know how bundling works and understand that you have to spend some time reading up on bundling. However, they would likely not be content with a bundling presentation that doesn't include any machine learning. It may not matter for the class, but if I were doing this for a paying customer I would do a quicker survey of bundling methods just to show the types of methods I am looking into, and then show preliminary results of machine learning on the easiest or most common bundling technique. Showing preliminary results of machine learning doing a good job on the most basic technique is a better milestone that would convince the customer to pay for the rest of the period of performance. Ultimately in business the person you or I interact with has to convince their boss to give them money to fund the effort, and actual machine learning results are what your customer will need to convince their boss to keep the money flowing. Again, this may not matter for the class, but it is something I have learned in my work experience and may be useful to consider for the future.