I would reclassify the purchase data as either essential or nonessential. The idea would be to try and see If we can identify a difference in people paying on time and what the person is using a credit card to purchase – life essentials or entertainment / going out. I would label groceries and utilities as essentials and dining out and entertainment as nonessential. One we have the Essential Vs Nonessential spend we can plot the data and use the classifier to help us identify the likelihood of someone paying. Because we are looking at a dollar to dollars comparison we don’t need to.

The KNN classifier would be helpful in using purchase behavior to determine a person’s likeliness to pay a bill. Are people who are paying utilities on a credit card doing it out of desperation and won’t be able to pay on time or is it the opposite. We would want to choose a large K in order to make sure one-person purchase behavior does not have to much of an effect on the outcome.