Acme Telephonica (AT) is a mobile phone operator that has customers across every state of the U.S.A.

AT struggles with customer **churn prediction**—customers leaving AT for other mobile phone operators.

AT hired us to take a new approach to reducing customer churn.

This case study is to develop a machine learning solution to this business problem.

AT did not approach us with a well-specified machine learning problem. Instead, the company approached us with a business problem—reducing customer churn.

Our first goal is to convert this business problem into a machine learning problem and develop a concrete solution.

To evaluate the available data, we have the data definitions.

Feature	Description
BILLAMOUNTCHANGEPCT	The percent by which the customer's bill has changed from last month to this month
CALLMINUTESCHANGEPCT	The percent by which the call minutes used by the customer has changed from last month to this month
AVGBILL	The average monthly bill amount
AVGRECURRINGCHARGE	The average monthly recurring charge paid by the customer
AVGDROPPEDCALLS	The average number of customer calls dropped each month
PEAKRATIOCHANGEPCT	The percent by which the customer's peak calls to off-peak calls ratio has changed from last month to this month
AVGRECEIVEDMINS	The average number of calls received each month by the customer
AVGMINS	The average number of call minutes used by the customer each month
AVGOVERBUNDLEMINS	The average number of out-of-bundle minutes used by the customer each month
AVGROAMCALLS	The average number of roaming calls made by the customer each month
PEAKOFFPEAKRATIO	The ratio between peak and off peak calls made by the customer this month
NEWFREQUENTNUMBERS	How many new numbers the customer is frequently calling this month?

Feature	Description
CUSTOMERCARECALLS	The number of customer care calls made by the customer last month
NUMRETENTIONCALLS	The number of times the customer has been called by the retention team
NUMRETENTIONOFFERS	The number of retention offers the customer has accepted
AGE	The customer's age
CREDITRATING	The customer's credit rating
INCOME	The customer's income level
LIFETIME	The number of months the customer has been with AT
OCCUPATION	The customer's occupation
REGIONTYPE	The type of region the customer lives in
HANDSETPRICE	The price of the customer's current handset
HANDSETAGE	The age of the customer's current handset
NUMHANDSETS	The number of handsets the customer has had in the past 3 years
SMARTPHONE	Is the customer's current handset a smart phone?
CHURN	The target feature

Midterm analysis

- Q1. Understand the data. Note any issues with missing or damaged data and how you handle it. Comment on whether normalization of features would be helpful. One paragraph.
- Q2. Develop an analysis of the churn data using two different classifiers. First, justify why you chose each classifier. Provide a description of your classifiers. More than one paragraph.
- Q3. Compare the performance of the classifiers. Determine which is preferable and justify your decision. More than one paragraph.
- Extra credit. The company now tells you new information. Adjust your analysis for the case where giving service to a churner costs the company $700 and excluding a customer who is not a churner costs $100, while other cases are 0. Now optimize for the least cost. Describe how this changed your classifier performance.