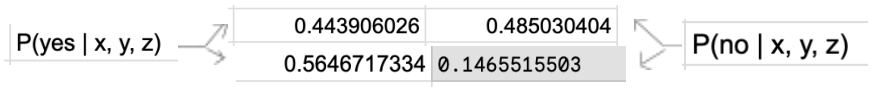
Justin Milner
V00906688
March 2024

SENG 474 Assignment 3

Question 1:

See attached completed titanic spreadsheet for full calculations and explanations. For a quick overview of the probabilities, the top row in the diagram below represents (2nd, child, male) and the bottom row represents (2nd, adult, female).



For question 1 I counted the number of passengers for each feature (2nd, adult, child, male, female, survived, not-survived). Then determined the probability of each feature given they survived and again the probabilities given they did not-survive. To determine the probability of surviving given the input features or not-surviving given the input features I then applied the Naive Bayes formula using the aforementioned probabilities.

Question 3:

a) I determined the prediction rating for Michael for the movie “You, Me, and Dupree” will be 1.875. I attached a spreadsheet containing all my calculations leading to this value. To determine the rating I first computed the mean rating of each individual user. Then I computed the PCC between michael and each individual user, which is the cosine similarity while subtracting the mean from each feature. I then took the summation of the correlation coefficients multiplied by the respective users’ rating for “you, me and dupree” and divided by the absolute value summation of all the correlation coefficients to arrive at the prediction rating for Michael.

b) I determined the user bias of lisa rose after the first iteration to be -0.10898807. I calculated the overall average rating of all the movies and used that along with the learning rate to update the user bias of lisa rose by iterating over each of her movie ratings.