Title:

Predicting Message Paths and Determining Conversation Intent from Live Chat Transcripts

Abstract:

We are given a data set from Comm100 and with this data set we predict the intent of conversations as well as cluster each message. We do this based on key word counts and other engineered features. We use the resulting clusters as states to build a Markov Chain transition matrix which models paths of messages between clusters, especially examining the visits to undesirable message states (angry users). We also implement an algorithmic (non-random) decision tree to categorically classify the conversation. We defined eight key topic categories based on the FAQ page from the Comm100 website.