Determining If A Coffee Chat Has Been Scheduled
Hong Kim, Minsu Kim
Advisor: Chris Callison-Burch

Motivation
- FreeForCoffee
  Personal side project. Service that pairs users to grab coffee. Used by over 1000 Penn students and creates up to 100 matches a week.
  Would like to know how many actually schedule coffee.
- Other services such as Tinder, OkCupid, and Airbnb could also share the same problem.

Goal
- Determining if two people scheduled to grab coffee given a text message exchange.

Data
- 1359 matches.
- 13,103 text messages.

Results / Evaluation
- Decision tree trained with three boolean features:
  ① Location exists?
  ② Agreement follows time and location?
  ③ Cancel follows last agreement?

  • Labeled data (gold standard)
    Accuracy: 95.6%
  • Predicted label data
    Accuracy: 86.0%
  • Decision tree trained with whitelist-based label prediction data is a simple and effective solution to this problem.

Labeling
- Built a web interface to label anonymized data.
- Four labels: Time, Location, Agreement, and Cancel.

Input: Text Message (SMS) Exchange
Decision Tree Classifier
Output: Scheduled YES/NO