Midpoint Report

CS 510 - Explorations in Data Science

Yiming Zhang

Paper Topic Updated

In recent years, the practice of data mining has become a popular computing technology. Data mining is a means to summarize useful information by analyzing big sets of data. In this project, I will research the definition, history and the process of data mining and how data mining affects our data privacy. I dive deep into how Google collects user's data. A case study of digital privacy issues amid the COVID-19 outbreak will be explored.

Research Plan Updated

Data privacy is a very broad topic. In this project, the main topics I will explore include: what is data mining and how data mining influences people's daily life in both good ways and bad. In detail, I will research how big companies like Google collect your personal information and how they use it to benefit you or potentially harm you. In addition, I will do a case study of COVID-19's impact on data privacy, protection and security. I take the advice from Kristin to remove the data collection study from Facebook and Amazon and focus on the case study of COVID-19's impact on data privacy.

Team Structure Updated

There is only one team member in this team which is me. No update for this part.

Progress with respect to Project Milestones:

Regarding the project milestones, I am proceeding right on schedule. I have already finished part 1 - 3 of the milestones (Introduction to data mining and its history, the good and bad side of data mining in terms of data privacy and how Google collects your personal data). No changes have been made concerning the rest of the milestones. I will focus on the case study of COVID-19's impact on data privacy and writing up the rest of the paper in the second part of the term.

Midpoint Meeting - Date & Time:

Mon Jul 21, 2020 1:40 pm – 1:50 pm (PDT)