Project Proposal

Math189R - Mathematics of Big Data 1

Nico Espinosa Dice

February 17, 2020

1 Project Overview

For my project, I will focus my area of research on biometric authentication. Biometric authentication is a field of security that uses biological characteristics of individuals to verify identity. My goal is to develop a machine learning model that can authenticate individuals based on biometric data.

2 Data

I will likely be using biometric data generated by smartphones, primarily through the accelerometers, gyroscopes, and magnetometers embedded within them. I have found several datasets on Kaggle that supply the necessary data I will need for my project. I have listed links to potential datasets (and their corresponding Kaggle competitions, if applicable) below:

- Human Activity Recognition Using Smartphones Data Set
- Accelerometer Biometric Competition
- Hand Tremor Dataset for Biometric Recognition.
- Activity Recognition

3 Methods

I have found academic papers that can help guide my approach towards this problem. I have listed them below:

- New Behavioral Biometric Features for Continuous Authentication of Smartphone Users
- Evaluation of Hand Micromovement Features for Continuous Authentication of Smartphone Users During Typing
- Human Activity Recognition Using Smartphones Data Set
- A Data Reduction Scheme for Active Authentication of Legitimate Smartphone Owner Using Informative Apps Ranking[1]

I believe that this project offers an opportunity to apply the probabilistic perspective that we have been learning in class and through our reading. (Bayes' theorem will be especially applicable). I am currently not set on the algorithm that I will use, as I have significantly more research to do before deciding that, but I know that this project offers the opportunity to apply neural networks, gradient vector machines, KNN, and more.

4 Team

Currently, I do not have any partners for this project. While I would enjoy being on a team, most other students seem to already have decided on another project. That said, I am untroubled by working on this project alone.

[1] This paper uses individuals' app use, rather than physical sensor measurements, for verification. While this is slightly off from my project's intention, it may nonetheless be useful.