Ilya Shafir
UCSD AI/ML Bootcamp - March 2024 Cohort
ishafir@gmail.com

13.4.4 Capstone Project Ideas
24ᵗʰ April 2024

(A) FRAUD (ANOMALY) DETECTION FOR PHARMACY INSPECTION REPORTS

Statewide pharmacy inspectors frequently need to analyze thousands of prescription and prescription fill records for any one pharmacy to validate the pharmacy is in compliance with local regulations. This is a time consuming, cumbersome task, that may be offset by an ML system that can quickly spot anomalous data for the inspector to review. The data tends to come in simple spreadsheet format so it's relatively easy to clean and then send through the ML model. The main concern for the project would be to gather real data and anonymize sufficiently so to preserve any privacy concerns. Other options is to potentially obtain free sample data sources available online.

(B) FRAUD DETECTION FOR GENERALIZED PATIENT HEALTHCARE DATA

If option (A) is not feasible, then going for a more generalized approach with more freely available data in this space.

(C) ML MODEL FOR CASUAL YET NOT SIMPLY SOLVED GAME

Nothing as ambitious as AlphaGo but this would be a very interesting path to create an ML model to play against a player in a game that has more complexity than chess, but perhaps easier to play. Not sure if a game like Lines and Boxes in a sufficiently large matrix may be a contender but something in that ilk, perhaps.

(D) PREDICTING ELECTIONS USING SOCIAL MEDIA SENTIMENT ANALYSIS

Twitter and other social media platforms are full of opinions about politics and candidates in government elections. There are publicly available datasets that can be analyzed. An alternative would be to gather data directly from Twitter in order to perform sentiment

analysis on previous election cycles and see if there's a true correlation to the election results.