Capstone Project 2: Initial Project Ideas

Topic 1:- Netflix Movie Recommendation

Problem Statement:-
Netflix's bread and butter is making sure people are greeted with the movie they like, when they login. Netflix uses a movie recommendation system which predicts whether someone will enjoy a movie based on how much they liked or disliked other movies.
Our aim here would be to build a recommendation system that would give a 10% better results than the existing recommendation system.

Dataset:-
https://www.kaggle.com/netflix-inc/netflix-prize-data

Solution/ Proposal:
We would explore different types of approaches and ML models, maybe build an ensemble model and check, if we can make better predictions here.

Topic 2: Stock Prices Prediction

Problem Statement:-
Great fortune has been made and lost in Stock markets. Betting on the right stocks can lead to enormous returns. Similarly, stocks with low returns can be disastrous for ones' financial health.
We would want to predict stock prices by using the past/ historical prices (of stocks) as input.

Dataset:-
https://www.kaggle.com/borismarjanovic/price-volume-data-for-all-us-stocks-etfs
Full historical daily price and volume data for all US-based stocks and ETFs trading on the NYSE, NASDAQ, and NYSE MKT.

Solution/ Proposal:
Using the ML models we build a recommendation engine, that would recommend us the stocks that have better chances of appreciation.

Topic 3: Covid19

Problem Statement:-

Dataset:-
https://github.com/CSSEGISandData/COVID-19
This dataset has daily level information on the number of affected cases, deaths, and recovery from 2019 novel coronavirus. This is a time-series data and so the number of cases on any given day is the cumulative number.
The data is available from 22 Jan 2020.

Solution/ Proposal:

Topic 4: Airbnb Recommendation

Problem Statement:-
Airbnb is a platform which joins the guest (travellers) to the hosts, presenting a more unique, personalized way of experiencing the world.
To maintain a smooth experience, we need to ensure that prices of the places on offer, are regulated and projected properly.

Dataset:-
http://insideairbnb.com/get-the-data.html
https://www.kaggle.com/dgomonov/new-york-city-airbnb-open-data
Airbnb has made the dataset public, which contains information pertaining to user booking, such as dates, reviews etc.

Solution/ Proposal:
We would implement a recommendation model, that would recommend the hotel prices based on the metrics available with us.