

Capstone Two Project Ideas

Idea one: The West Nile virus Prediction. For this project, we need to predict whether the West Nile virus is present for specific regions in Chicago. Every year the public health workers in Chicago set up mosquito traps in the city to collect mosquitos. Mosquitos are then tested for the presence of the West Nile virus. In this project we have the historical weather data of Chicago as well as the historical mosquitos test data for each region in the city. We need to predict the presence of the virus in the future based on the data we have. We will also identify the most important factors that influence the presence of the virus and make suggestions to state legislators. Our dataset comes from Kaggle:

<https://www.kaggle.com/competitions/predict-west-nile-virus/data>

Ideal two: Road accident analysis and prediction. For this project, we are given comprehensive information about road accidents in France from 2005 to 2016, such as the weather condition, road condition, vehicle information. We need to predict the severity of the accident and identify the key factors that contribute to the most severe accidents. We can then make suggestions on precautions to reduce the number of severe accidents. Our dataset comes from Kaggle:

<https://www.kaggle.com/datasets/ahmedlhlou/accidents-in-france-from-2005-to-2016>

Idea Three: customer purchase behavior analysis and prediction. For this project, we are given household level transactions over two years from a group of 2,500 households who are frequent shoppers at a retailer. We also have many customers' demographic information as well as direct marketing contact history are included. Based on these information, we can study customer purchase behavior, whether coupons and campaigns actually improve customers' purchase. Our database comes from Kaggle: <https://www.kaggle.com/datasets/frtgnn/dunnhumby-the-complete-journey>