Project 4

West Nile Virus in Chicago

DSI 01 DIP PREM DOE

Agenda

1. Introduction
2. Exploratory Data Analysis
    a. trap
    b. temperature
    c. spray
3. Model
    a. baseline
    b. regression + classification
    c. final model (excluding weather data + explanation)
4. Evaluation
    a. confusion metrics
5. Cost Benefit Analysis

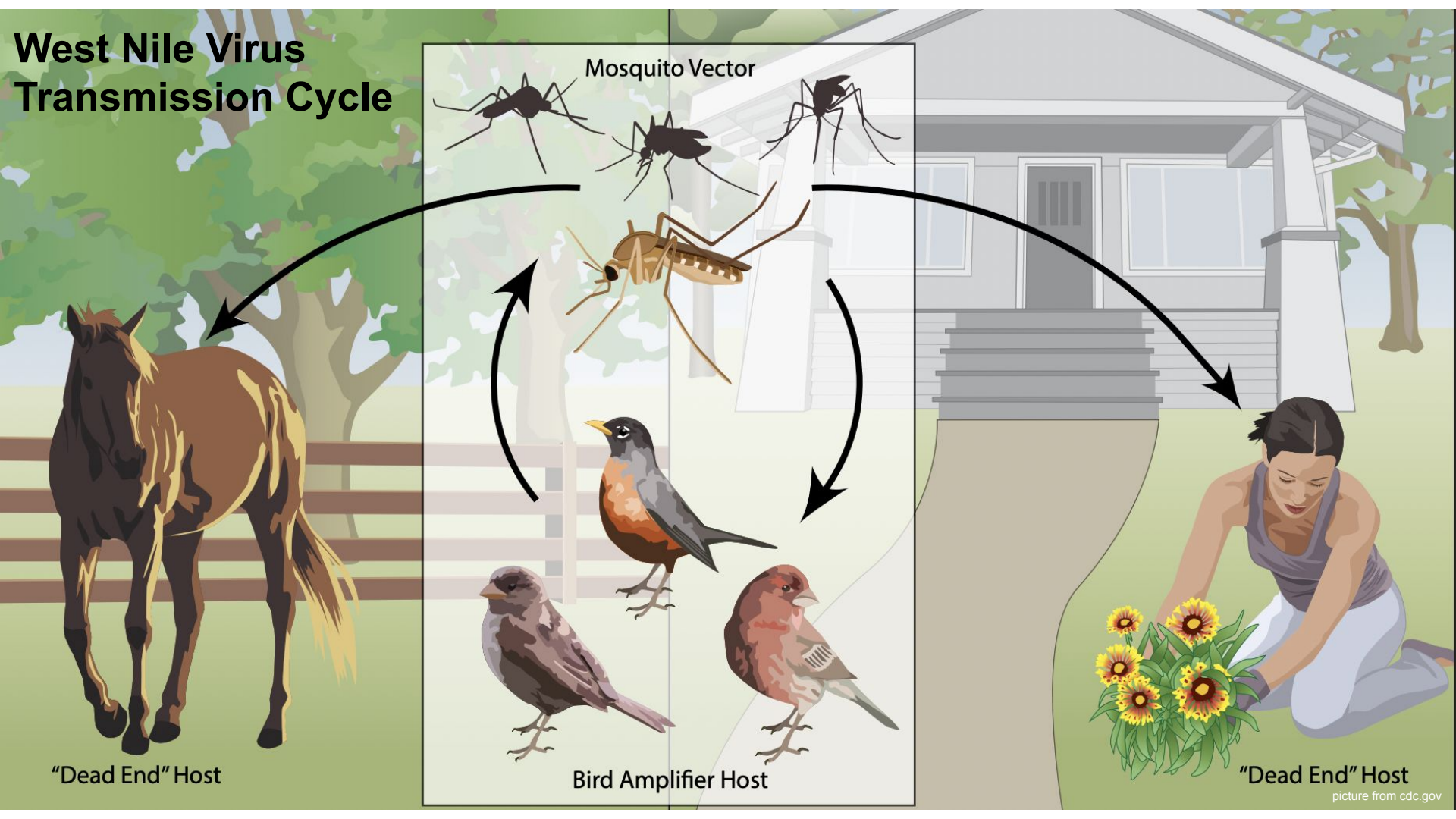
Introduction

What is West Nile Virus?

West Nile virus (WNV) is the leading cause of mosquito-borne disease in the continental United States. It is most commonly spread to people by the bite of an infected mosquito. Cases of WNV occur during mosquito season, which starts in the summer and continues through fall. There are no vaccines to prevent or medications to treat WNV in people. Fortunately, most people infected with WNV do not feel sick. About 1 in 5 people who are infected develop a fever and other symptoms. About 1 out of 150 infected people develop a serious, sometimes fatal, illness. You can reduce your risk of WNV by using insect repellent and wearing long-sleeved shirts and long pants to prevent mosquito bites.

picture from cdc.gov

West Nile Virus Transmission Cycle

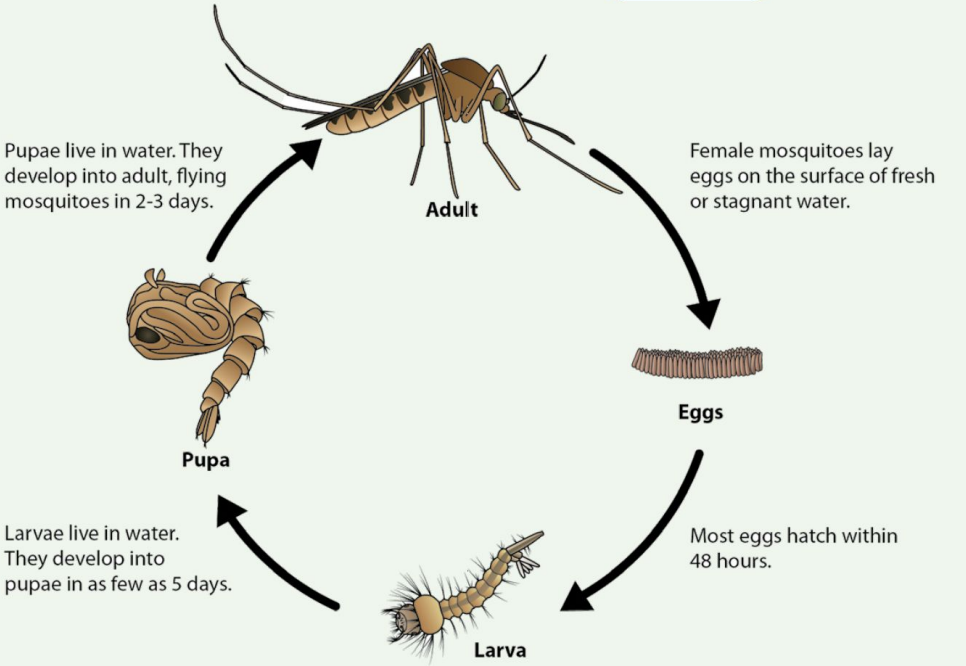


Mosquito Life Cycle

Culex pipiens, *Cx. quinquefasciatus*, and *Cx. tarsalis*

It takes about 7–10 days for an egg to develop into an adult mosquito.

U.S. Department of Health and Human Services
Centers for Disease Control and Prevention



Life stages of *Culex* species mosquitoes

Eggs

- Adult, female mosquitoes lay eggs on the surface of fresh or stagnant water. Water sources can include barrels, horse troughs, ornamental ponds, unmaintained swimming pools, puddles, creeks, ditches, and marshy areas.
- A female *Culex* mosquito lays eggs one at a time. Eggs stick together to form a raft of 100 to 300 eggs. The raft floats on the water.

Larva

- Larvae hatch from mosquito eggs and live in water.
- Larvae can be seen in the water. They are very active and are often called "wigglers."
- They feed on materials found in the water.
- Larvae shed their skin (molt) several times during this stage.

Pupa

- Pupae live in water. Pupae do not have external mouthparts and do not feed during this stage.
- An adult mosquito emerges from a pupa and flies away.

Adult

- Adult female mosquitoes bite people and animals. Mosquitoes need blood to produce eggs.
- After blood feeding, female mosquitoes look for water sources to lay eggs. Several days pass between feeding and looking for a place to lay eggs.
- *Culex* mosquitoes don't fly long distances, but have been known to fly up to 2 miles (3.2 km).
- Some *Culex* mosquitoes prefer to live near and bite birds. They bite people when other animals are not nearby.
- Because *Culex* bite animals and people, they live outdoors or near homes.

For more information on diseases spread by mosquitoes:

- www.cdc.gov/westnile
- www.cdc.gov/easternequineencephalitis
- www.cdc.gov/sle


Eggs stick together to form a raft.


Larvae in the water.


Pupae in the water.


Female mosquito **before** a blood meal.


Female mosquito **after** a blood meal.

picture from cdc.gov

West Nile Virus in Chicago

CHICAGO – The Chicago Department of Public Health (CDPH) is reporting 6 cases of West Nile virus infection this year among Chicago residents, following CDC confirmatory testing performed for the first case. All patients developed symptoms between August 15-September 5 and were hospitalized. No deaths have been reported. The age range is 43-75 years (median 57), 3 are female, and 4 are White and 2 are Hispanic. One individual reports traveling out of state during their potential exposure period. The individuals reside in the North, West and East regions of the City. For the 2019 season, CDPH reported 6 human cases, including 1 death.

West Nile virus is transmitted through the bite of a *Culex pipiens* mosquito, commonly called a house mosquito, which has picked up the virus by feeding on an infected bird. Common symptoms include fever, nausea, headache and muscle aches. Symptoms may last from a few days to a few weeks. However, four out of five people infected with West Nile virus will not show any symptoms. In rare cases, severe illness including meningitis, or even death, can occur. People older than 60 and individuals with weakened immune systems are at higher risk for severe illness from West Nile virus and need to remain vigilant.

Each year CDPH conducts a comprehensive mosquito surveillance program, which includes placing larvicide in catch basins to limit the number of mosquitoes that can carry the virus, and trapping mosquitoes throughout the city and testing them for West Nile virus. By using data, the City is able to most efficiently target high-risk areas for the virus and keep residents safe.

news copy from https://www.chicago.gov/city/en/depts/cdph/provdrs/healthy_communities/news/2020/september/first-human-cases-of-west-nile-virus-in-chicago-for-2020.html

Problem Statement

Predict when and where different species of mosquitoes will test positive for West Nile virus. A more accurate method of predicting outbreaks of West Nile virus in mosquitoes will help the City of Chicago and CPHD more efficiently and effectively allocate resources towards preventing transmission of this potentially deadly virus.

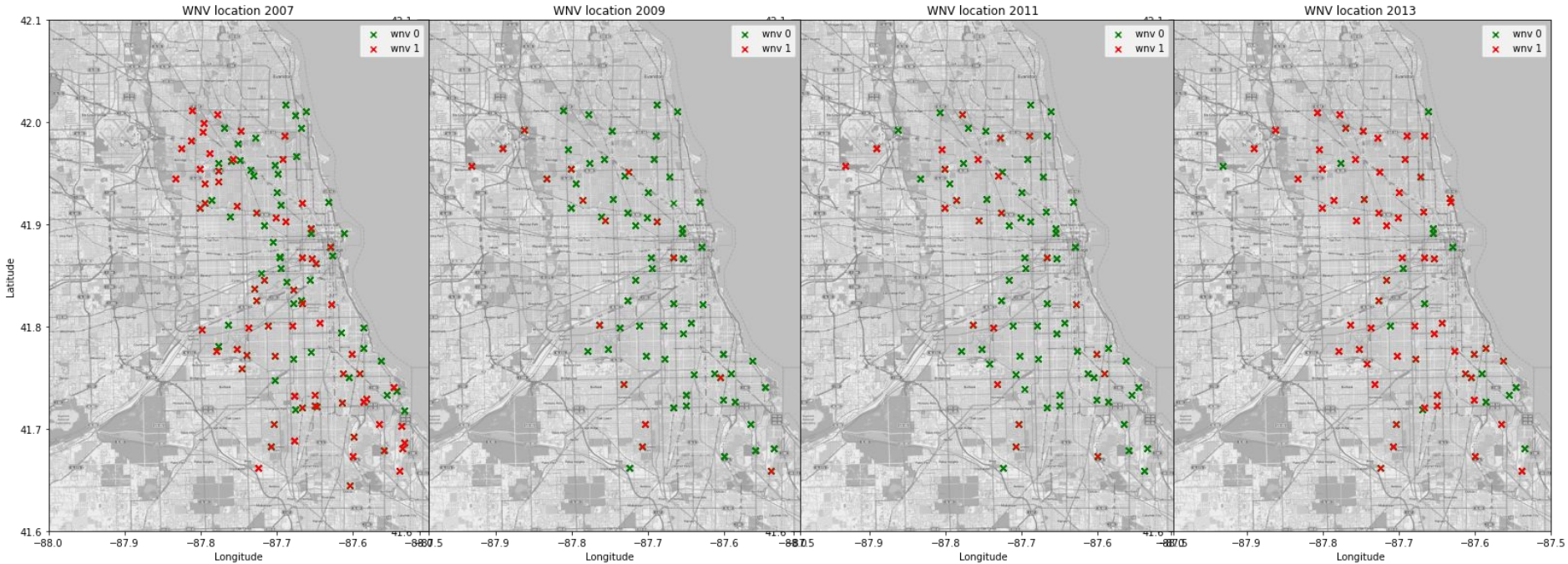
Exploratory Data Analysis

Data

1. Mapdata : Chicago map
2. Trap : 2007, 2009, 2011, 2013 trap locations, mosquito in the trap, wnv present
3. Weather : from 2 stations, May2007 - Oct 2014
4. Spray : Spay date and location(latitude, longitude)

TRAPS vs WNV present

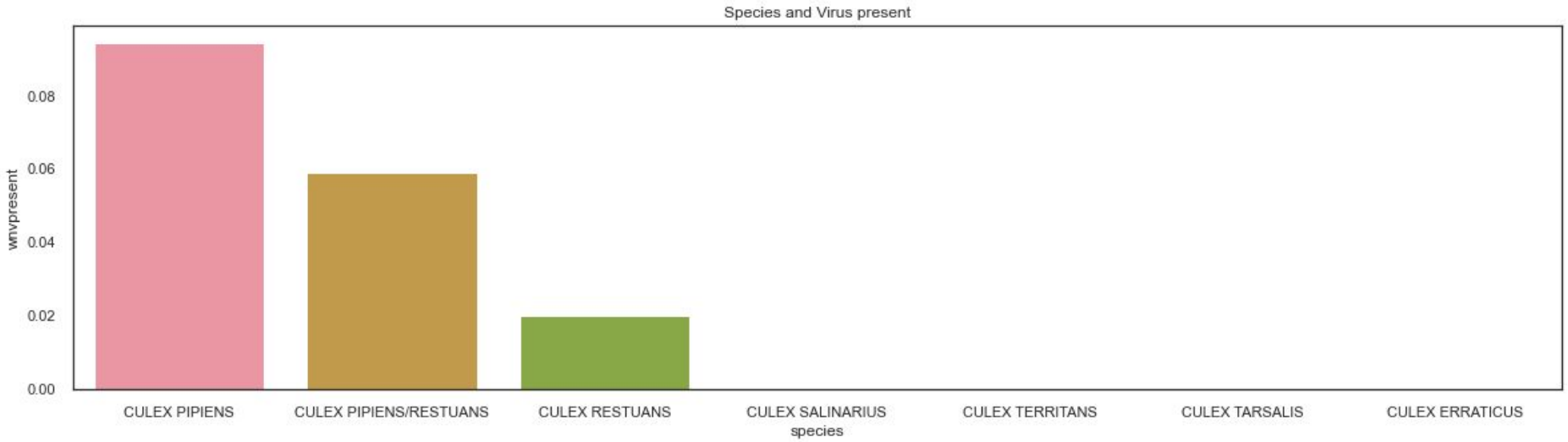
2007 2009 2011 2013



x Trap without WNV present x Trap with WNV present

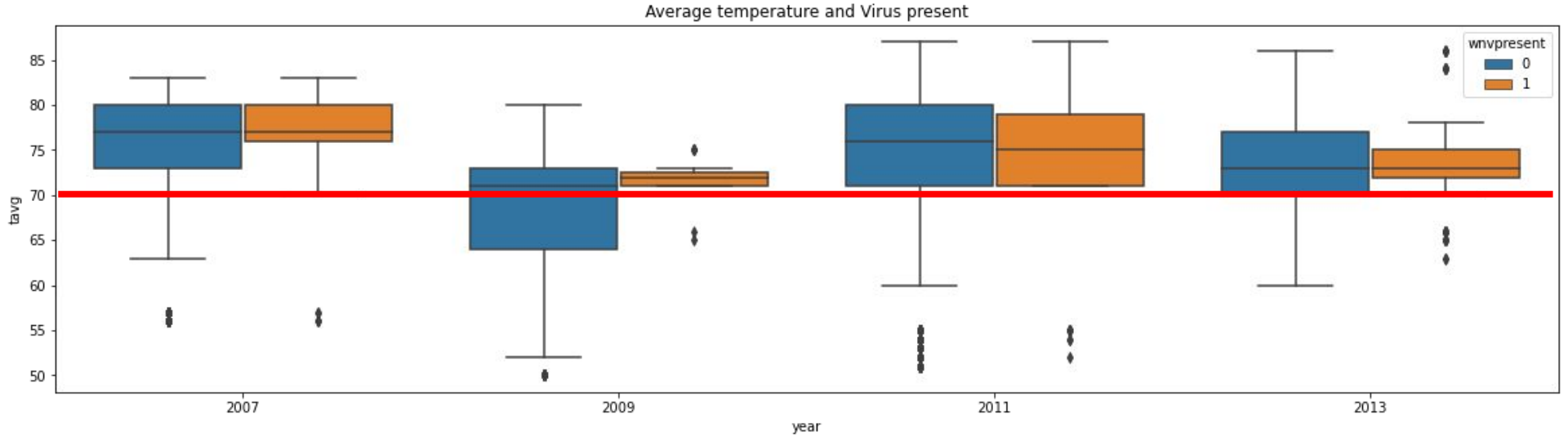
West Nile Virus presence is mostly dominated in the year 2007 and 2013.

Mosquito species that be carrier of the WNV



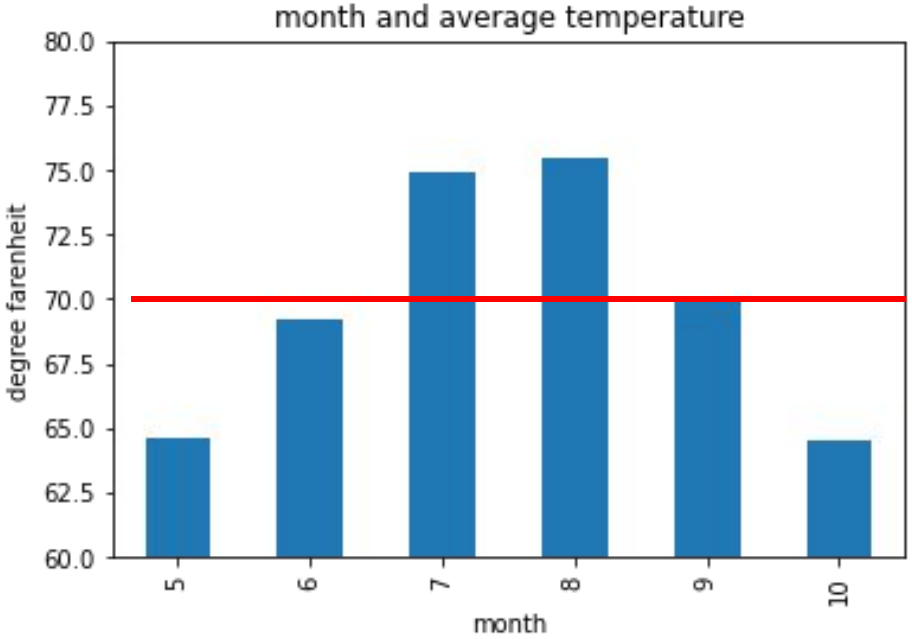
CULEX PIPINES, CULEX PIPINES/RESTUANS, CULEX RESTUANS

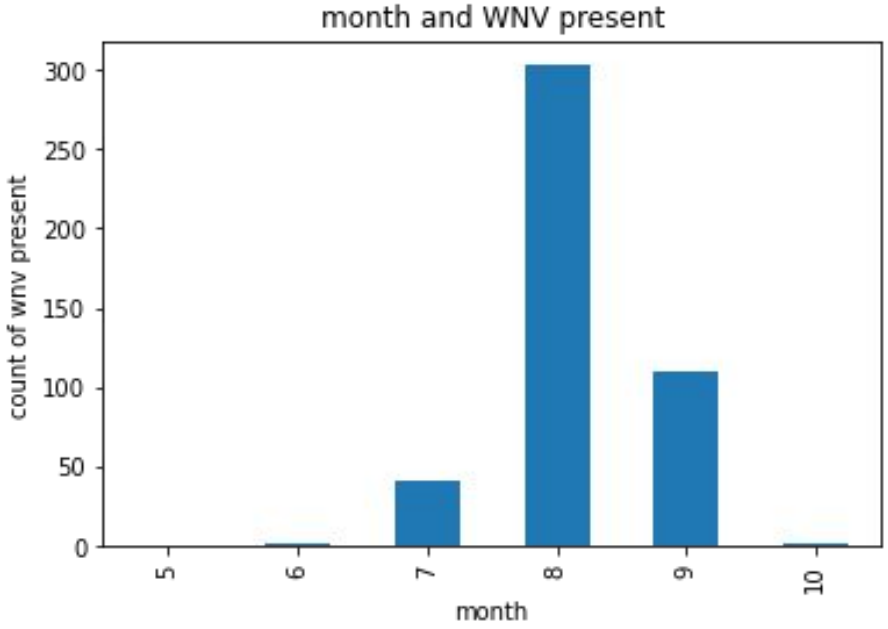
Temperature and WNV present



There are very few cases that WNV present at lower than 70 degree fahrenheit.
Year 2009 has the lowest average temperature and the lowest virus spread.
We think there are some correlation between temperature and WNV present.

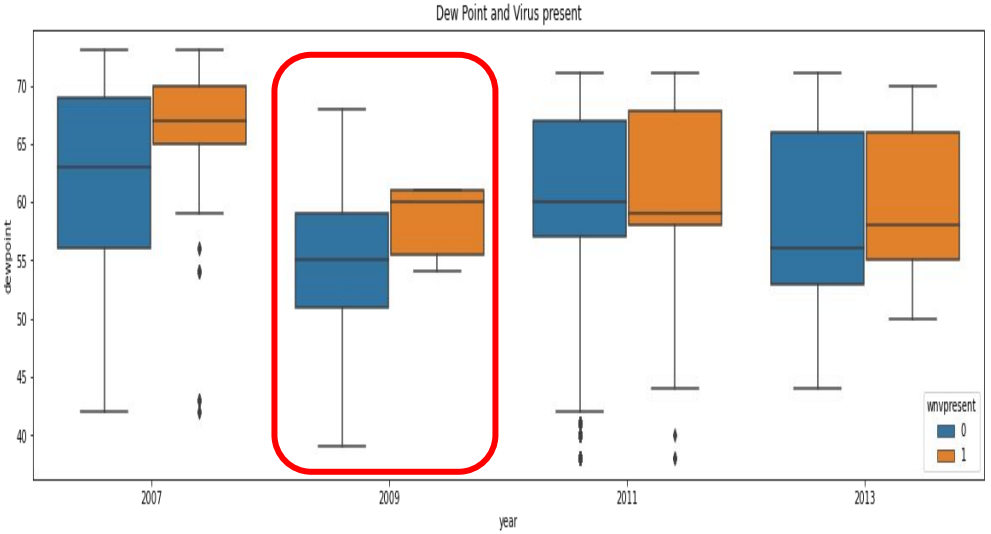
Temperature and WNV present

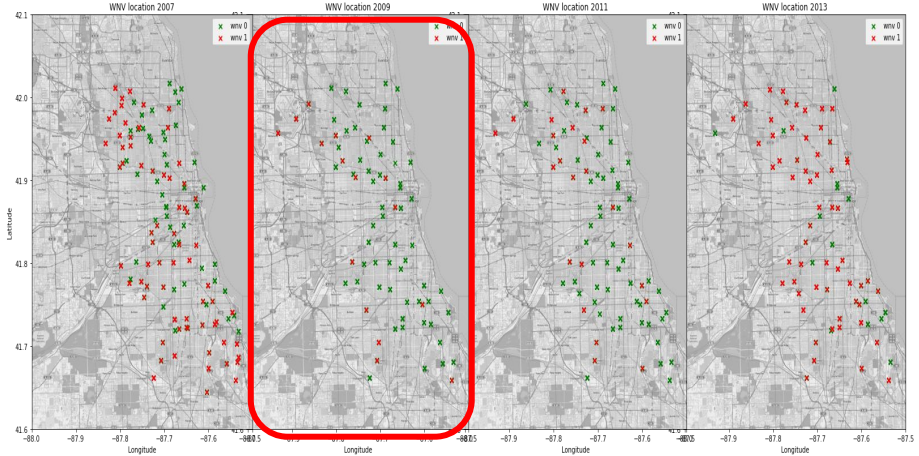




Above graph confidently confirm relation between temperature and WNV present. Temperature is peak at the same time of WNV present.

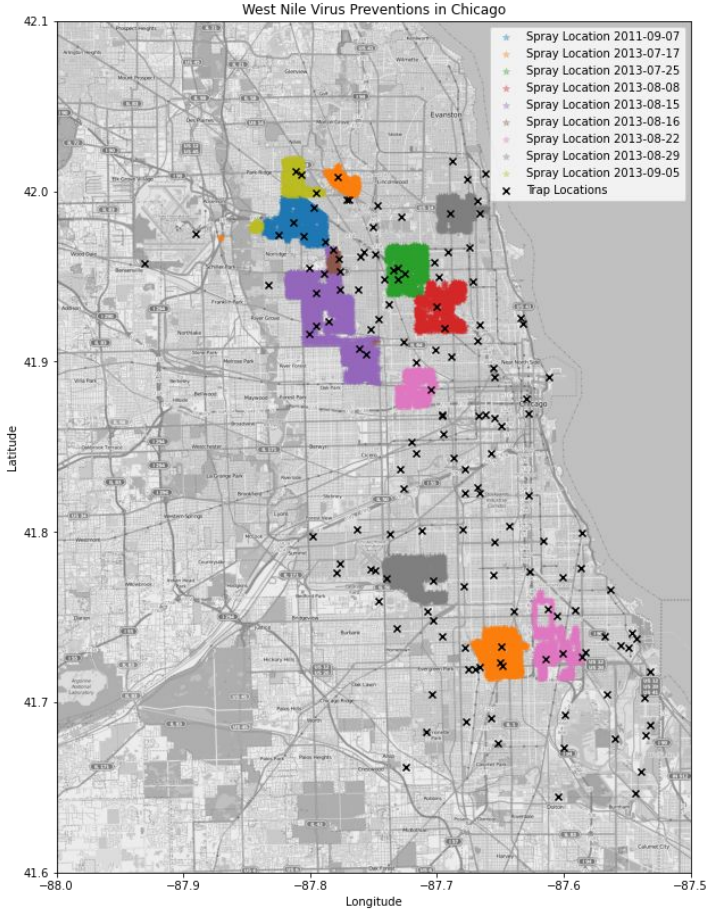
Dew point





From above graphs, an increase in mean dew point temperature tend to associate with WNV present. Mean dew point temperature in 2009 is the lowest, in that year WNV present is also the lowest(compare to 2007, 2011,2013).

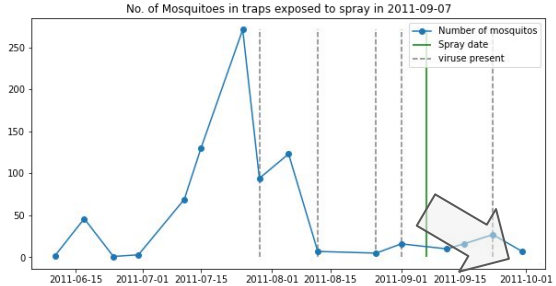
Pesticide spray date and location

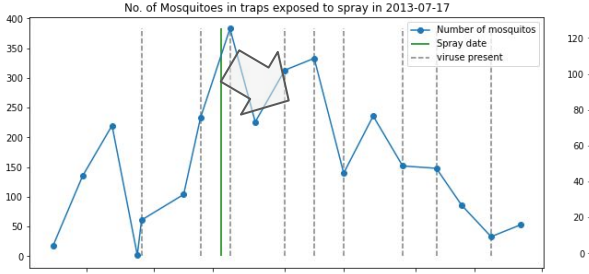


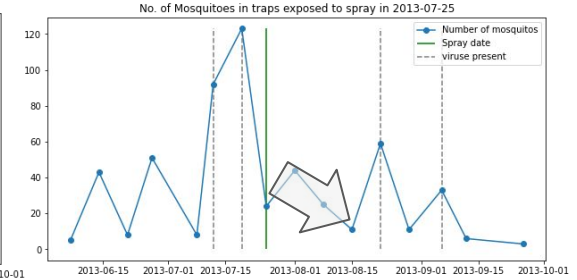
In dataset, they did spray several times to control and reduce the number of mosquitoes that can spread viruses.

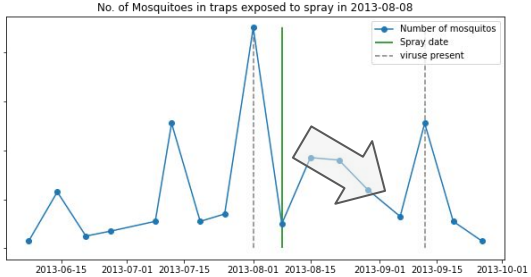
Date and area of spray show at the left map.

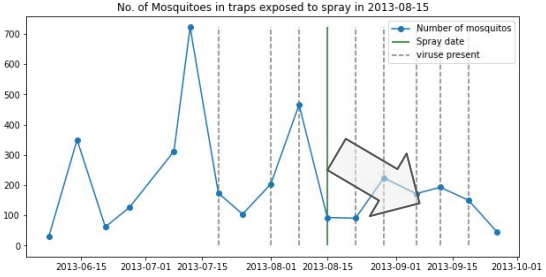
Pesticide Spray and Number of Mosquito

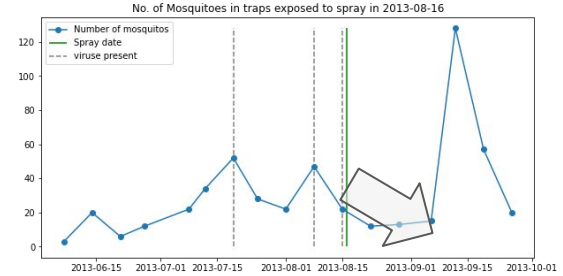


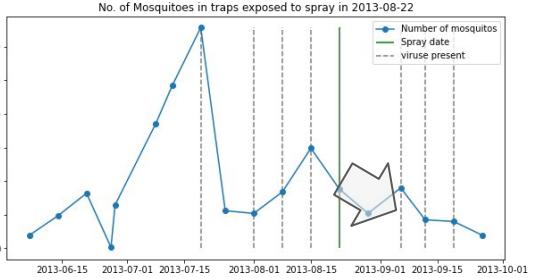


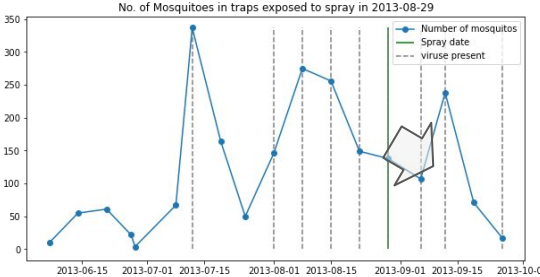


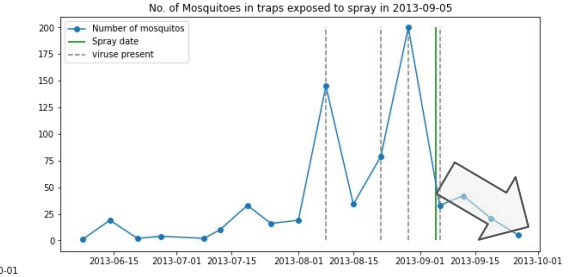








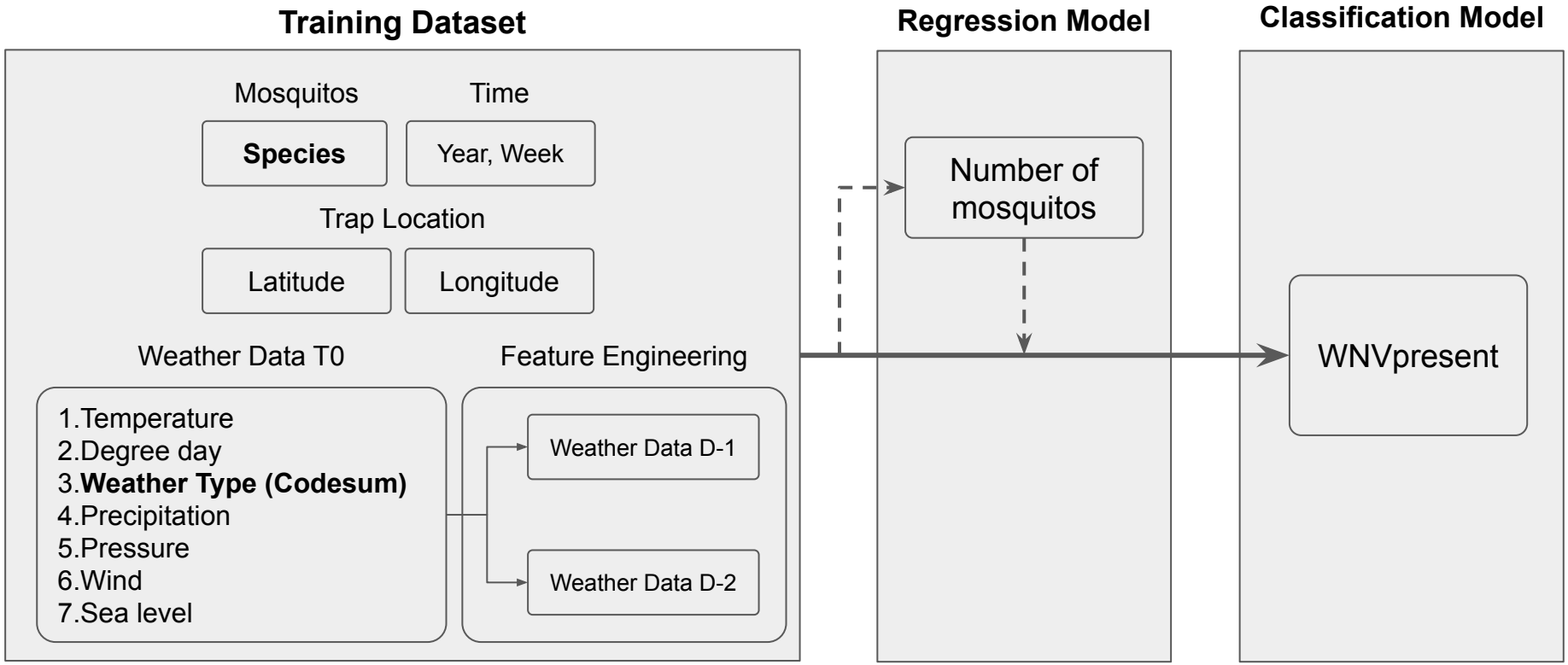




The result of spray did not eradicate the presence of virus but only to reduce the number of mosquitos in the area. And the result did not last long.

MODEL

Modeling Process



*Note: Species and Codesum are converted into dummy variables

Model Selection

Regression Model

1. Linear Regression
    a. Ridge
    b. Lasso
    c. ElasticNet
2. KNeighborsRegressor
3. Decision Tree-Based Regressor
    a. RandomForestRegressor
    b. ExtraTreesRegresspr
    c. AdaBoostRegressor
    d. GradientBoostingRegressor
    e. BaggingRegressor
4. SVR

Scoring : Accuracy

Classification Model

1. LogisticRegression
2. KNeighborsClassifier
3. Decision Tree-Based Classifier
    a. DecisionTreeClassifier
    RandomForestClassifier
    b. AdaBoostClassifier
    c. GradientBoostingClassifier
4. SVC

Scoring : ROC-AUC

Final Model: RandomForest Regressor + GradientBoost Classifier

Features

1. Average temperature
2. Dewpoint
3. Species - Dummy
4. Week
5. Year
6. NumMosquitos (Predicted)

Regression

Random Forest Regressor

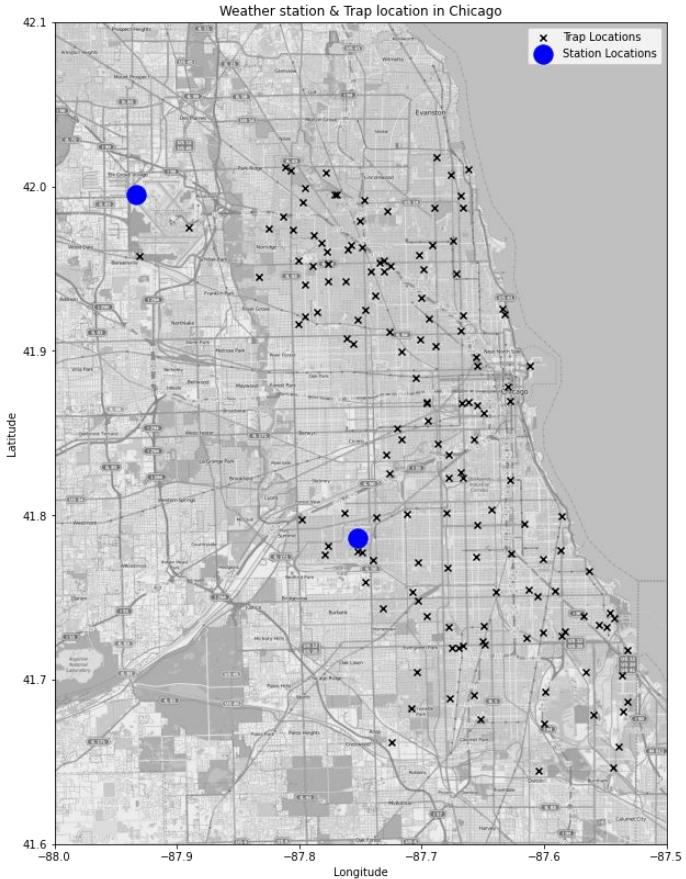
R2 Score
Training : 0.9
Validating : -0.39
RMSE: 67.81

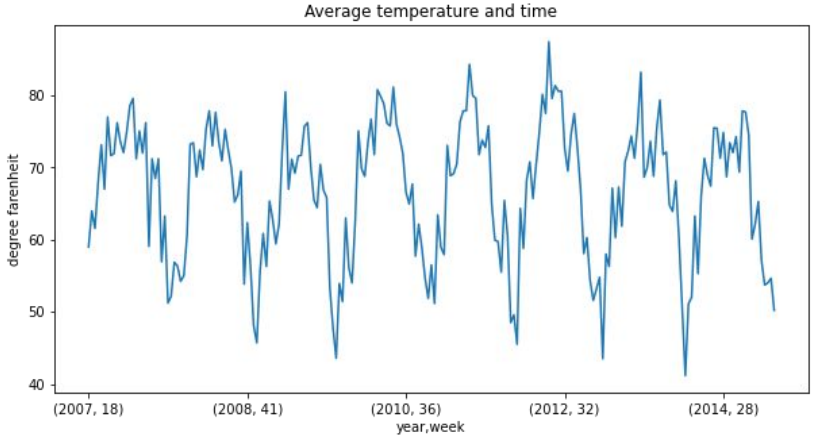
Classifier

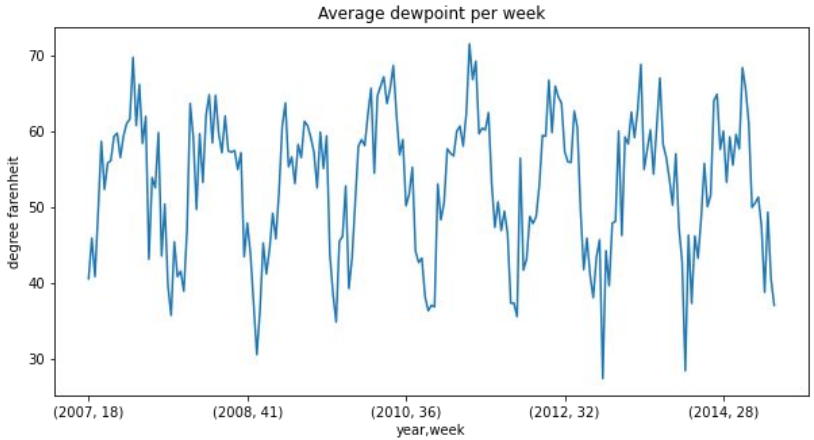
Gradient Boosting Classifier

ROC_AUC Score:
Validating : 0.88768

Week vs Weather Data

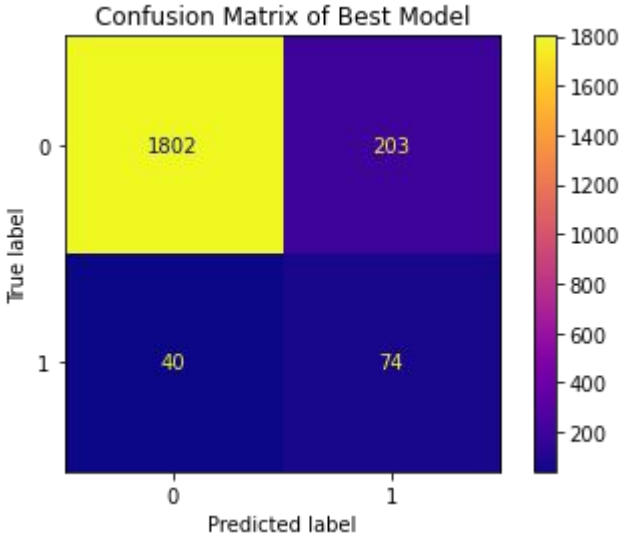


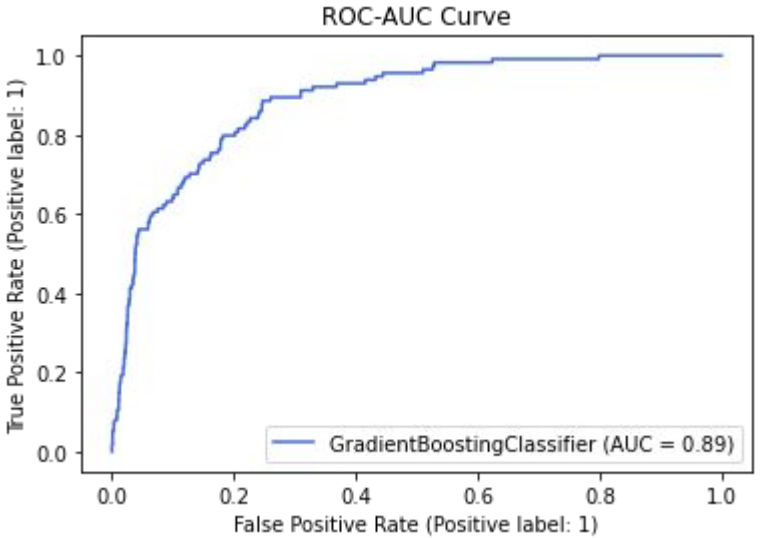




EVALUATION

Evaluation - Best Model





We need to optimize for Sensitivity (False Negative)

Misclassification: 11%

1 - West Nile Virus present
0 - West Nile Virus not present

Ways to optimize for Sensitivity

1. One of the limitation for our data is the training dataset that is given on a two year interval basis. If data was on a year to year basis, more accurate patterns might come show up.
2. Find more precise indicators in our data through more feature engineering, for example, combination of weather patterns.
3. Since this virus is a transmitted through a cycle, the root cause is not only mosquitos, but other animals as well. Hence, researching on other animals that transmit the disease might provide a better indication on how it affects mosquitos.

Evaluation on Test Dataset

#	△pub	Team Name	Notebook	Team Members	Score	Entries	Last
1	▲1	Cardal			0.85990	150	6y
2	▲3	Dmitry & Leustagos			0.85757	170	6y
3	▲1	nhlx5haze			0.85619	170	6y
4	▲3	The Iron Curtain			0.85345	165	6y
5	▼4	no one			0.85253	159	6y
6	▼3	Victor			0.85114	166	6y
7	▼1	Syowen			0.84933	145	6y
8	—	May the Force be with us			0.84574	281	6y
9	▲3	Silogram			0.84218	126	6y
10	▲7	_Sineksavar_			0.84090	94	6y

Submission and Description	Private Score	Public Score	Use for Final Score
model_3.csv 17 minutes ago by Tanupong Rattanasawatesun add submission details	0.76536	0.77753	☐

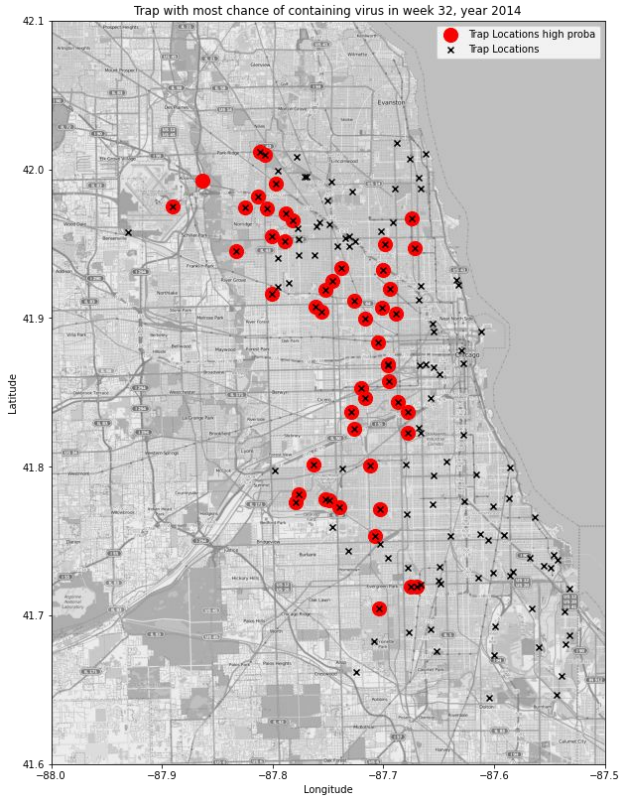
Cost - Benefit Analysis

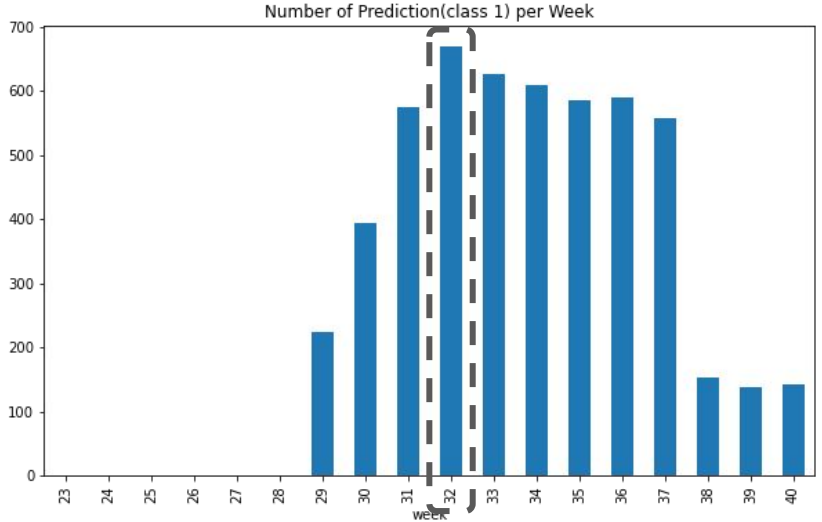
Criteria for spray recommendation

Traps that has high number of mosquitos
and also high probability of
West Nile Virus present

High crowded area

Selecting trap with highest chance





Costs

Not spraying:

- On average, the initial cost of a person infected with the virus is about $23,000.
- There is an average of 150 infections reported to CDC per year from 1999 to 2014 in Illinois.
- **This means that even if only 3 people are affected in Chicago, not spraying would also cost more than spraying.**

Spraying:

- Spraying once a month
- Each month costs around $100 to spray ½ acres
- High temperature months from June to September (4 months)
- Since we want to spray to reduce False Negative, that would mean spraying at all locations.
- **This means each year it would cost the city around $60,000 to spray.**

Next steps

- Tracking mosquito population growth and incorporating it into our model will greatly improve precision
- Evaluate alternative tactics, such as clearing dead birds
- Investigate the effectiveness of spraying insecticide with complete spray data
- Other methods put in place by the city also seems effective enough.
    a. Recommendations on repellents
    b. Drain water
    c. Repair windows and doors
    d. Wear protective clothing
    e. Report dead birds, high grass and standing water

Thank You