

## ABSTRACT

In this study, we use multiple data analytics techniques to analyze and predict illegal border crossings in the United States along the Canadian and Mexican borders. The motivation of this study is to show how using descriptive and predictive analytics could help making strategic decisions for reducing cross-border crime and illegal immigration. Our methodology compares the historic and current data to predict which variables have the most effect on illegal border crossing and analyze how illegal border crossing has changed over time in different locations and seasons. We posit our preliminary research might be of interest to the U.S. Immigration and Customs Enforcement (ICE) agency for a more strategic planning and resource allocation.

## INTRODUCTION

According to the border crossing data, the number of undocumented migrants that cross the United States border is estimated to be 10 million. With the problems progressively estimated to increase, controlling the movement of people and goods has become even more challenging and poses a greater threat to American society. So, it has become essential for border patrol agencies to rethink their border protection strategies. Below is the graph of increasing US Border Patrol budget, but it does not show an effective result in reducing the number of illegal border crossing.

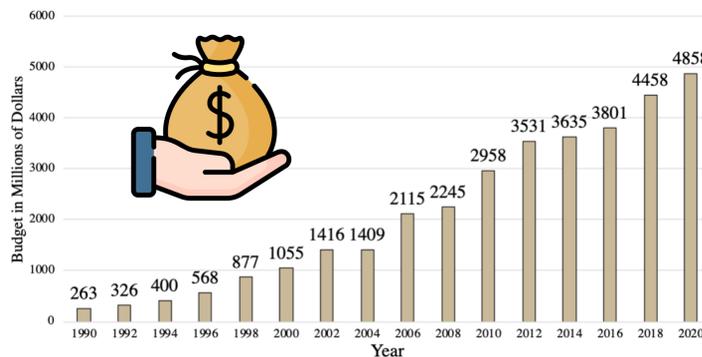


Fig 1 . U.S. Border Patrol Budget  
Source: <https://www.americanimmigrationcouncil.org>

## METHODOLOGY

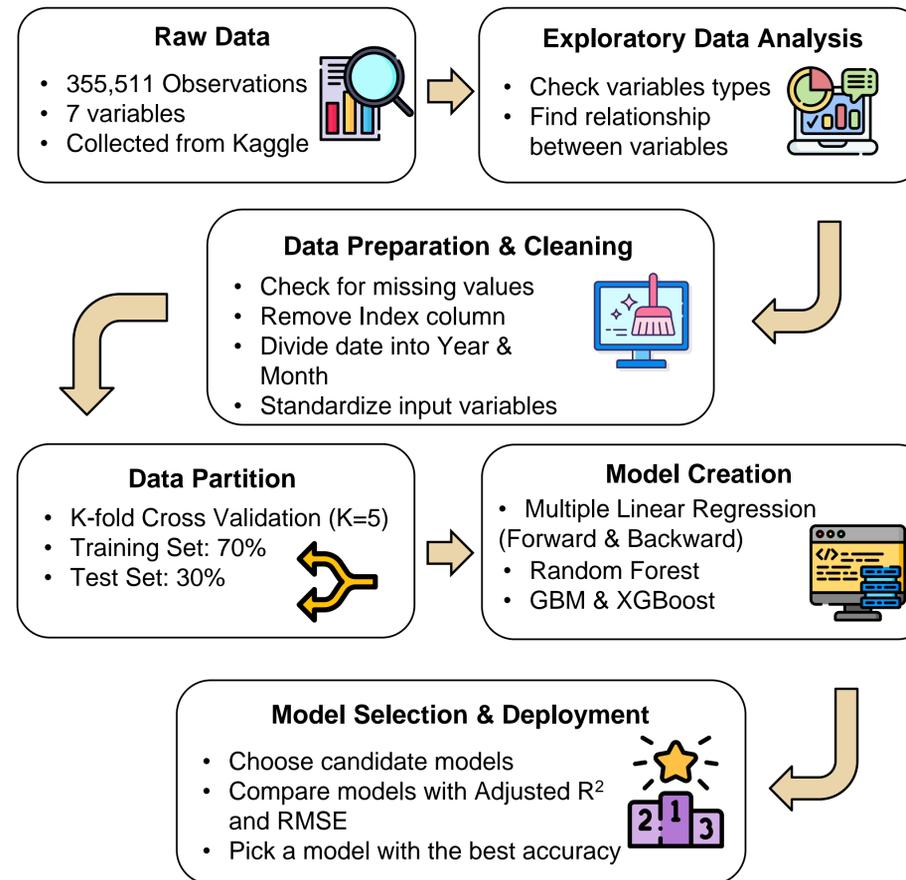


Fig. 3. Methodology

## STATISTICAL RESULTS

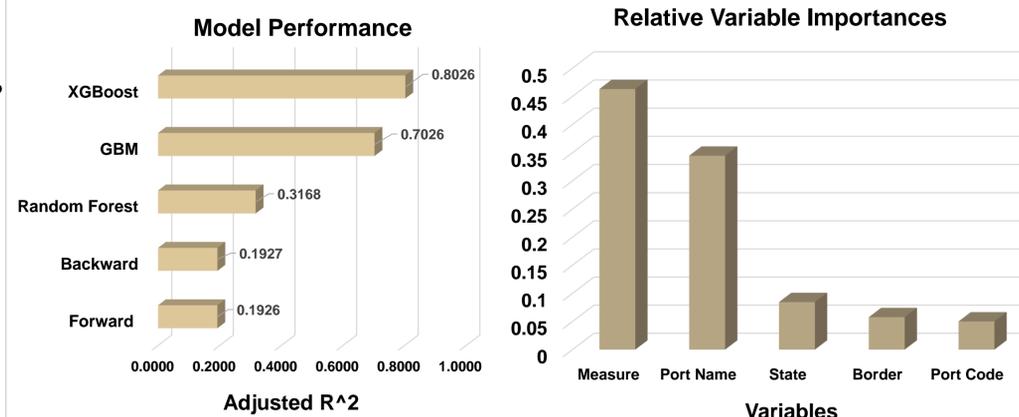


Fig. 4. Model Performance

Fig. 5. Relative Variable Importances

Performance Measure: R<sup>2</sup>  
Candidate Models: GBM & XGBoost

Best Model: XGBoost  
Adjusted R<sup>2</sup>: 0.8026

## EXPECTED BUSINESS IMPACT

- “Measure”, which is the method of transportation used when crossing the borders was the most relatively important variable.
- Illegal border crossings have been decreasing in the US-Canada borders, while increasing in the US-Mexico borders.

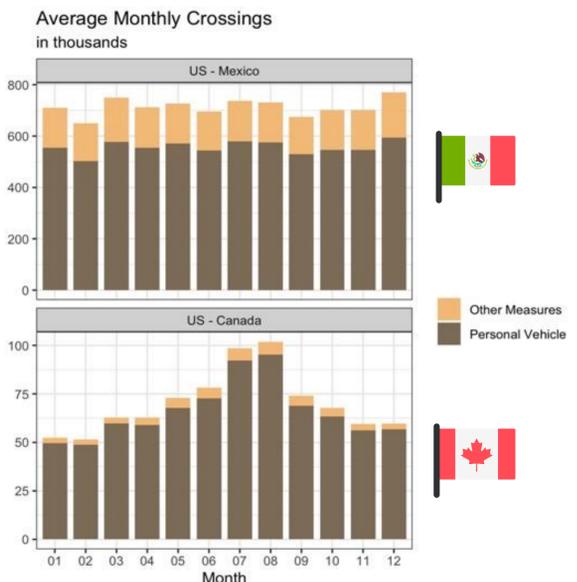


Fig. 6 Border Crossings by Month

- ICE could allocate more resources on strengthening security in checking different types of transportations (Measures) used during border crossing.
- ICE could use our model to predict illegal border crossings in different locations and seasons to better allocate their resource.

## CONCLUSIONS

- We analyzed several models to build the best model to predict illegal border crossings.
- “Measure” and “Port Names” are the variables that have the most effect on illegal border crossing.
- Illegal border crossing has been decreasing in the US-Canada borders and increasing in the US-Mexico borders.
- There has been a steady number of border crossings in the US-Mexico borders throughout the four seasons but increase in the US-Canada borders during summer.
- Our study can be improved in the future if we had more detailed datasets on crimes happened near borders.

## ACKNOWLEDGMENTS

We would like to thank Professor Matthew Lanham, Xinyu Wang, Theo Ginting, and our graduate student mentors Yujin Tao and Zhewei Tao for their guidance and support on this project.

### Checklist for ICE

- More resource in US-Mexico borders throughout the year
- Security strengthening during summer in US-Canada borders
- Check other types of vehicles like trucks and buses at the borders



### Research Objectives:

- Which **variables** have the most effect on illegal border crossing?
- How has illegal border crossing varied over time in different **locations** and **seasons**?

## LITERATURE REVIEW

The literature that we examined on illegal border crossing is summarize below. We found these studies do not investigate each individual variable that may impact illegal crossings. Also, they do not consider the spatial nor temporal changes based on season and location features. Therefore, our study is novel in that we use multiple models to analyze these variables and discuss some potential impacts of using our solution.

Study	Regression	SVR	SVN	SARIMA	SUR Modeling	Mathematical
Jaromi (2013)	√					
Lin (2015)		√	√	√		
Maoh (2016)					√	
Towers (2018)	√					
Mohammed (2018)						√

Fig. 2. Literature Review