# ACCIDENTALLY\_LATE

J. KIRK, S. MCNAIR, N. FOSTER, K. HOLMBERG, M. ONEAL

## SELECTED TOPIC

- -Traffic accidents 2018-2022
- -Impact of Covid: increase in accidents in Covid years vs. Pre-Covid years
- -Impact of weather: temperature, visibility, wind speed, precipitation, weather condition, sunrise vs sunset, year, and date

## REASON TOPIC SELECTED

This topic was chosen to see if a correlation exists between the severity of car accidents during the Covid years when compared to the severity of car accidents in pre-Covid years. In this context, "severity" refers to the delay a car accident causes in the traffic pattern. A longer delay would mean a more substantial car accident.

## WHO/WHAT MIGHT USE THE DATA

- -Weather Channels: utilize weather data to advise drivers of what to expect; delays, types of accidents, locations of increase of accidents
- -Insurance Companies: raise rates in high severity/frequent areas
- -Car rental agencies deciding price rates
- -Companies debating working from home and start times for work

#### TOP 5 STATES WITH THE MOST ACCIDENT OCCURRENCES

ACCIDENTS UPDATED	PreCOVID accidents	COVID accidents
CALIFORNIA (611,261)	California (125,281)	California (480,030)
FLORIDA (355,328)	Oregon (35,137)	Florida (342,487)
TEXAS (111,194)	Minnesota (17,267)	Texas (103,287)
OREGON (100,907)	Utah (13,115)	Virginia (89,056)
VIRGINIA (93,357)	Florida (9,460)	Pennsylvania (85,880)

### DESCRIPTION OF DATA SOURCE

- <u>US-Accidents: A Countrywide Traffic Accident Dataset Sobhan Moosavi</u> (<u>smoosavi.org</u>)
  - Dataset has 47 columns of information; we included the following columns for the purpose of the project: severity, start time, state, temperature, visibility, wind speed, precipitation, weather condition, sunrise vs sunset, year, and date
- cphalpert/census-regions: US Census Bureau Regions and Divisions by State (qithub.com)
  - US Census Bureau Regions and Divisions by State; 4 regions, 7 divisions

## QUESTIONS EXPECTED TO BE ANSWERED

What impact has Covid had on the severity of car accidents in the US?

Has weather played a role in the information?

#### Top 5 Weather Conditions with the most accident occurrences

	Accidents updated	PreCOVID accidents	COVID accidents
Fair	1044151	115198	927250
Cloudy	335487	41604	293072
Mostly Cloudy	288935	32253	253466
Partly Cloudy	196112	31529	163688
Light Rain	121335	18454	89987

### DATA EXPLORATION PHASE INFORMATION

According to the data, as of right now, the number of accidents went up during Covid years when compared to pre-Covid years

```
# Get count of unique values in the 'year' column from 'accidents_updated'
 2 print(accidents updated['year'].value counts())
2021
        1419863
2020
         575031
2019
         198861
2018
          21139
2017
          19517
          11009
Name: year, dtype: int64
 1 # Get count of unique values in the 'year' column from 'PreCOVID accidents'
    print(PreCOVID accidents['year'].value counts())
2019
        198861
         65722
        12458
Name: year, dtype: int64
    # Get count of unique values in the 'year' column from 'COVID accidents'
 2 print(COVID accidents['year'].value counts())
        1419863
2021
         508568
Name: year, dtype: int64
```

## **CONFUSION MATRIX**

$\alpha$	1
Sm	OTA
$\sim$ 111	$\cdot$

#### RFC

	Predicted long_delay	Predicted short_delay
Actual long_delay	7983	5416
Actual short_delay	14552	97049

	Predicted long_delay	Predicted short_delay
Actual long_delay	4068	9079
Actual short_delay	3363	108490

#### Naive\_RO

	Predicted long_delay	Predicted short_delay
Actual long_delay	8032	5367
Actual short_delay	14957	96644

#### Gradient BC

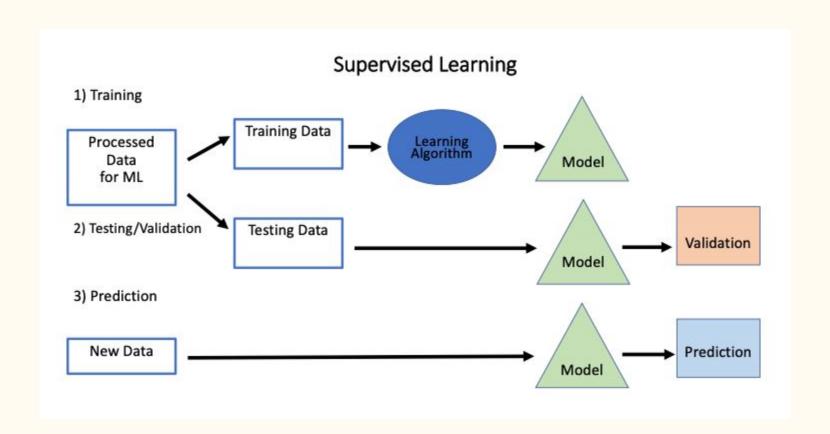
	Predicted long_delay	Predicted short_delay
Actual long_delay	1896	11566
Actual short_delay	1228	110310

#### SVM

	Predicted long_delay	Predicted short_delay
Actual long_delay	0	13399
Actual short_delay	0	111601

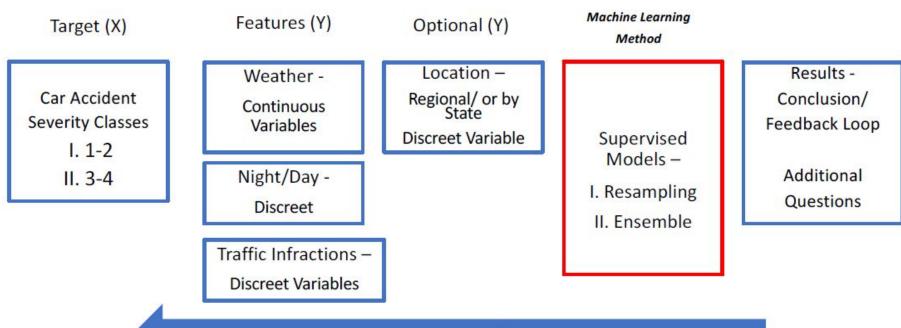
#### Smoteen

	Predicted high_risk	Predicted low_risk
Actual high_risk	7965	5497
Actual low_risk	14563	96975



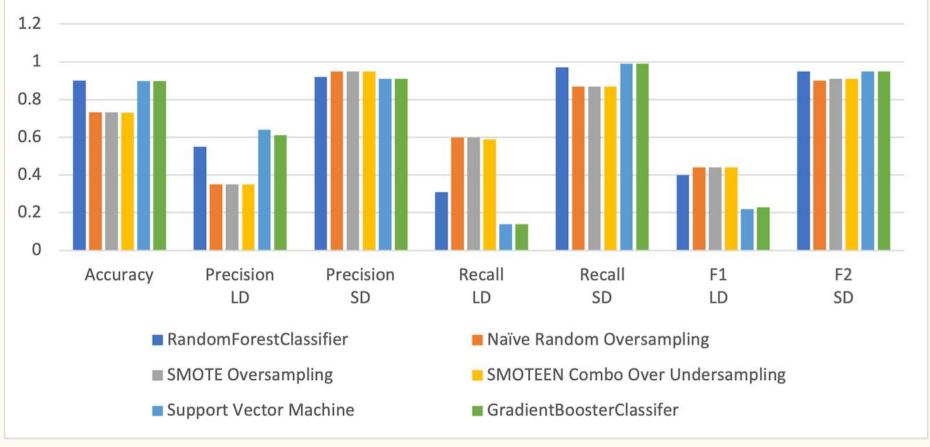
#### Machine Learning Diagram

Question: Were U.S. car accidents less severe during Covid year 2020-2021 compared to years prior?



REPEAT PROCESS/ REFINE





### ANALYSIS PHASE INFORMATION

#### $PRECOVID \rightarrow$

```
# Create a PreCovid Dataframe
# accidents_updated = pd.PreCOVID_accidents(date)
start_date = '2018-06-01'
end_date = '2020-02-29'
# Select DataFrame rows between two dates
mask = (accidents_updated['date'] > start_date) & (accidents_updated['date'] <= end_date)
PreCOVID_accidents = accidents_updated.loc[mask]
PreCOVID_accidents.head()</pre>
```

	severity	state_code	temperature_f	visibility_mi	wind_speed_mph	precipitation_in	weather_condition	sunrise_sunset	year	date
582028	4	NJ	89.0	10.0	5.0	0.0	Fair	Day	2019	2019-10-02
1295810	4	LA	54.0	10.0	5.0	0.0	Mostly Cloudy	Day	2019	2019-11-01
1537770	4	AZ	41.0	10.0	10.0	0.0	Fair	Night	2020	2020-02-16
1568027	4	TX	79.0	10.0	25.0	0.0	Light Rain / Windy	Day	2019	2019-06-04
1756843	4	VA	37.0	10.0	0.0	0.0	Fair	Night	2019	2019-10-19

	severity	state_code	temperature_f	visibility_mi	wind_speed_mph	precipitation_in	weather_condition	sunrise_sunset	year	date
224945	2	MA	42.0	10.0	12.0	0.0	Fair	Night	2021	2021-03-10
224946	2	CA	54.0	2.0	6.0	0.0	Fog	Night	2021	2021-07-30
224947	2	MD	79.0	10.0	9.0	0.0	Mostly Cloudy	Day	2021	2021-10-15
224948	2	WA	38.0	10.0	0.0	0.0	Cloudy	Day	2021	2021-12-21
224949	2	CA	52.0	10.0	9.0	0.0	Fair	Day	2021	2021-12-09

 $\leftarrow$  COVID

## **TECHNOLOGIES**



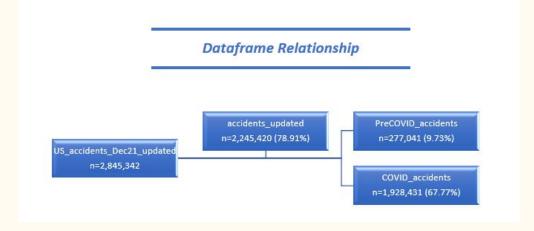
## **RESULTS**

More total accidents during Covid

Accidents updated	PreCOVID accidents	COVID accidents
2 (2,084,426)	2 (228,702)	2 (1,828,676)
3 (71,340)	3 (27,134)	4 (39,448)
4 (66,076)	4 (21,055)	3 (36,885)
1 (23,578)	1 (150)	1 (23,422)

## RECOMMENDATIONS FOR FUTURE ANALYSES

- Why is there > 1 million difference between PreCovid and Covid accidents?
- What impact did the COVID lockdown/quarantine have on the number and severity of car accidents in each state in the US?
- Impact of severity of individual states with lockdown vs. states without lockdown during Covid years.



## WHAT TO DO DIFFERENT NEXT TIME

• Change file structure at the beginning

### LINK TO VISUALS IN TABLEAU

https://public.tableau.com/app/profile/kurt.holmberg/viz/ProjectBook2\_16592207 802480/Story1?publish=yes