

SEARCH AND RESCUE

TIMEFRAMES

Technical Report

PROBLEM STATEMENT

Through analyzing this dataset for Search and Rescue missions, what are the key factors that indicate the timeframe to locate and rescue victims both in wilderness and non-wilderness environments?

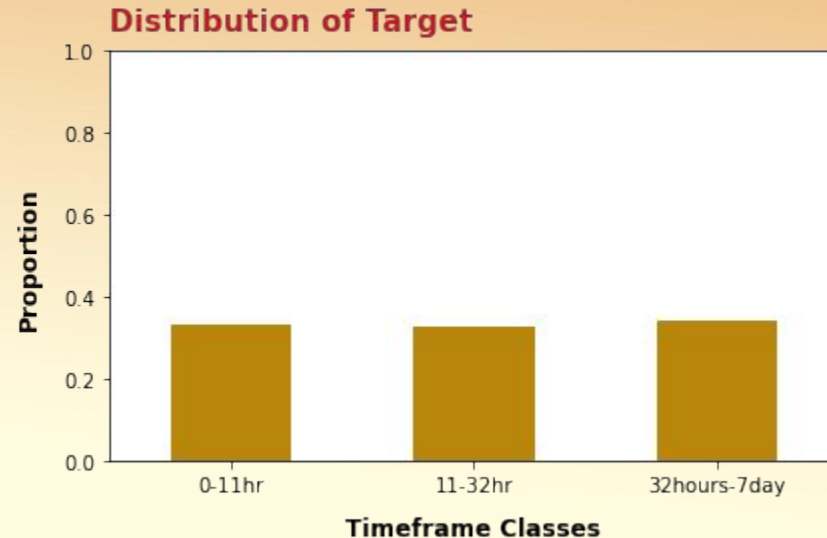


TARGET - SAR MISSION TIMEFRAMES

Classes: 0-11 hours | 11-32 hours | 32 hours - 7 days

- Normally distributed classes. Both classification models gained precision.
- Survival rates decrease when victims are in the elements especially overnight

Baseline: 35% for the majority class 32 hours - 7 days

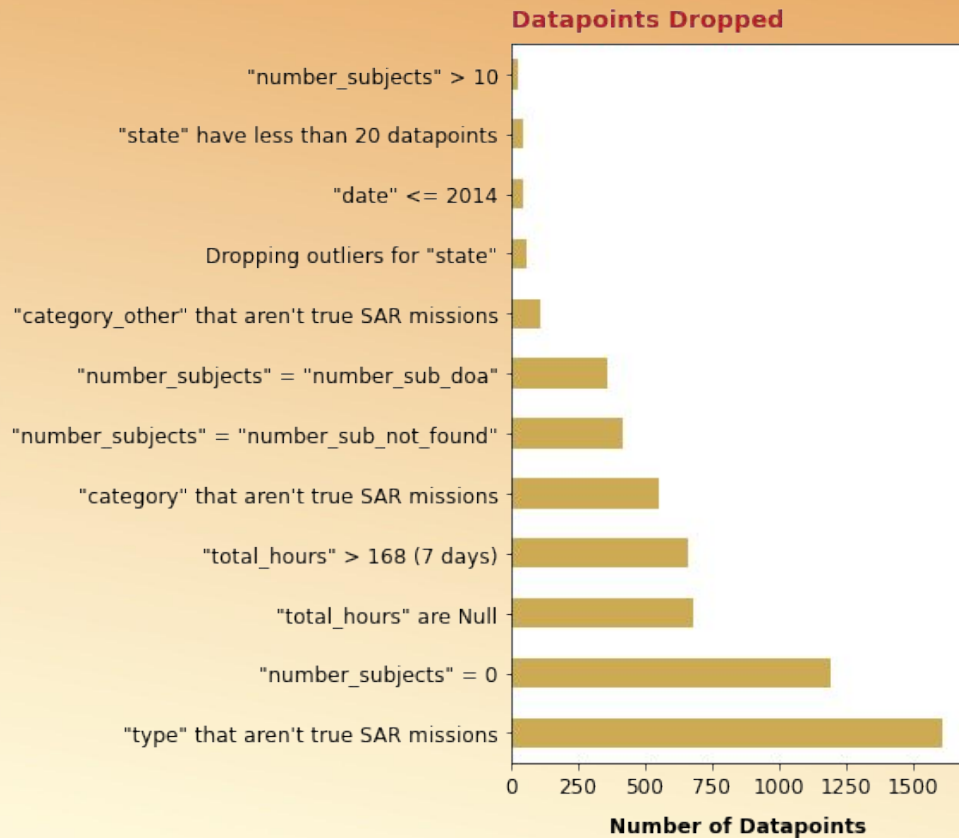


DATA CLEANING

Collected by Mountain Rescue Association

(<https://mra.org>)

- Live victim was being rescued
- 2014 - 2021
- 0-7 days
- Max 10 victims
- 13 States
- 78 SAR Teams
- 4 Area Types
- 10,000+ datapoints final



FEATURE ENGINEERING & IMPUTATION

State - imputed from Teams where State not null

Area Type - imputed nulls from other related variables

wilderness | urban/rural | water | interface

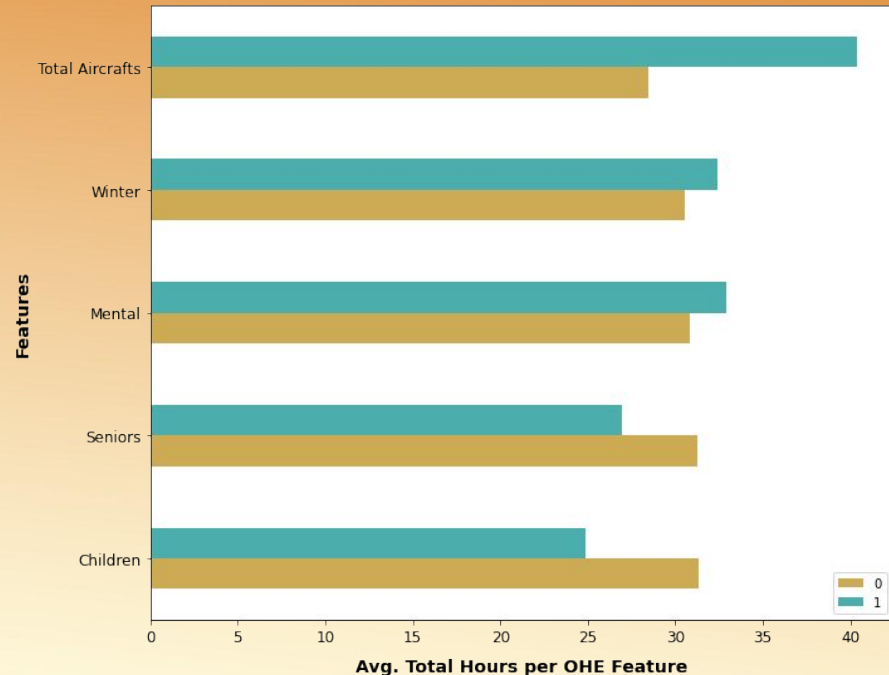
wilderness

Seniors - True where any subject had dementia

Winter - incident occurring in winter & snow conditions

Number Volunteers - Median

Engineered Features One-Hot-Encoded Avg. Total Hours



PREPROCESSING AND MODELING

Preprocessing

- All categorical variables were dummified
- All continuous variables were scaled using Standar Scaler

Modeling

- Tried SVM and a neural network
- Gridsearch was used for both Logistic Regression and Random Forest Models

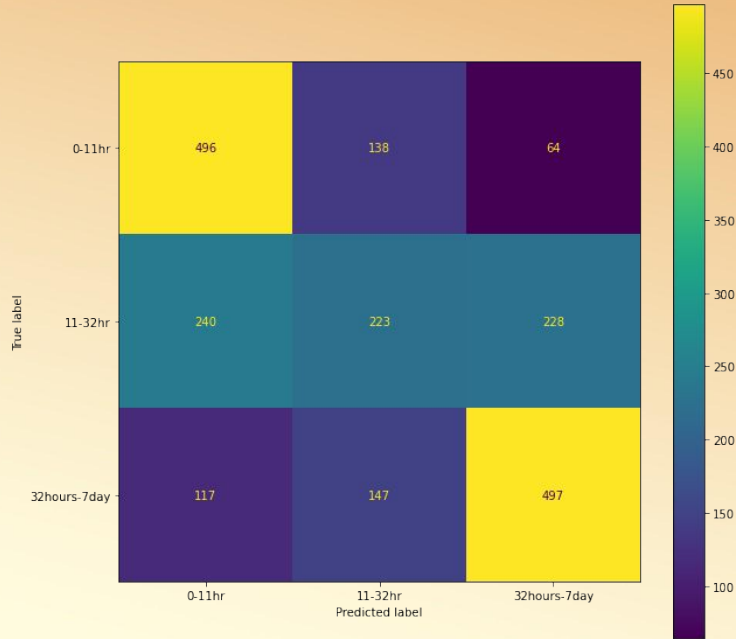


CLASSIFICATION MODELS

LOGISTIC REGRESSION

ACCURACY: 57% - 22% ABOVE BASELINE OF 35%

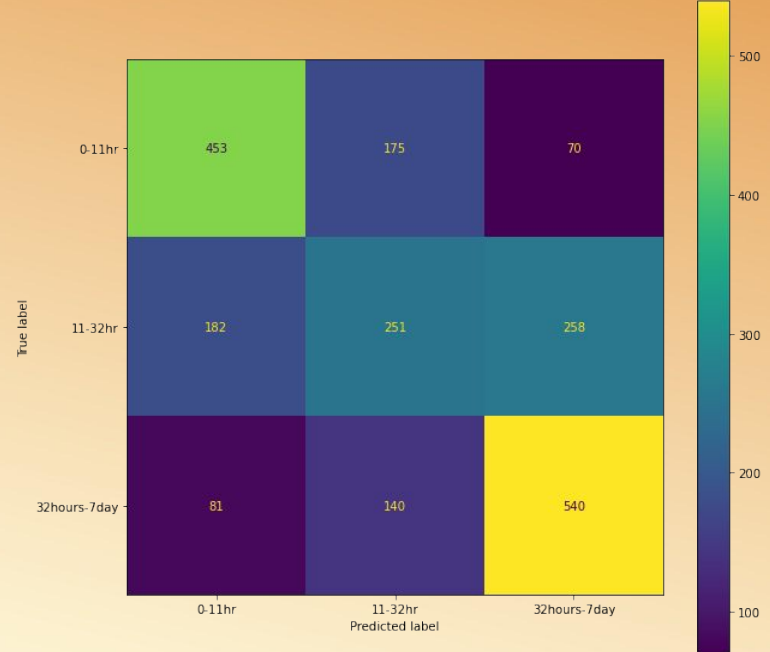
Precision for 2nd class $\frac{1}{2}$ of 1 and 3



RANDOM FOREST

ACCURACY: 58% - 23% ABOVE BASELINE OF 35%

Precision for 2nd class $\frac{2}{3}$ of 1 and 3



PRODUCTION MODEL

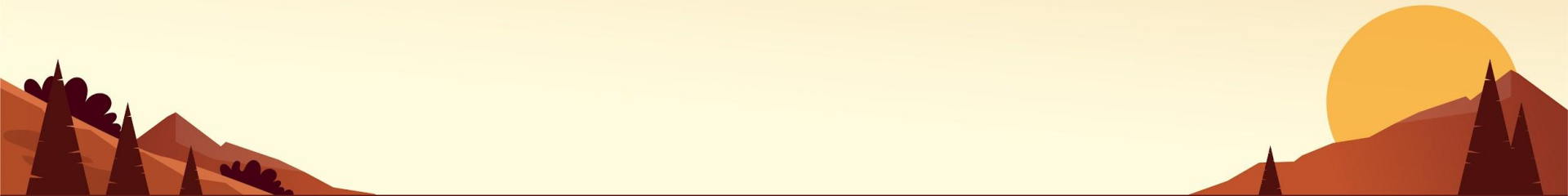
Logistic Regression with Lasso

- Most useful in understanding factors indicating each class

Random Forest Top Important Feature: Number Volunteers

Coefficients:

0-11 hours		11-32 hours		32 hours - 7 days	
Pennsylvania	29.85	New Jersey	0.49	New Mexico	1.85
Colorado	0.83	Nevada	-0.21	Washington	1.33
Number Volunteers	-0.74	Utah	-0.19	Number Volunteers	1.29



CONCLUSIONS & RECOMMENDATIONS

Production Model - Logistic Regression

This data set is very messy and could have more potential for gaining insight into SAR mission timelines. With cleaner data and more of it, machine learning models could more accurately predict the timelines for these incidents. If injuries were life threatening, models could predict the likely time the patient has before passing away and guide first responders processes and give them a better chance to save lives.

- Standardized data process
- Get more SAR teams to contribute data
- Allocate public funds & require standard procedures



THANKS! QUESTIONS



Andy Deemer