Avert Disaster:

Safety Modeling for Military Sealift Command (MSC) Ships


ENS Alexandros Psichas
2LT Charlson Ro

Advised By:
Dr. Allison Chang
Dr. Chelsea Curran
Dr. Mike Matthews
Mr. Jordan Levine

Sponsored By:
MIT Lincoln Laboratory,
Group 43,
Air Traffic Control Systems

Problem Statement

Problem Description: The Military Sealift Command (MSC) provides maritime logistical support such as fuel and supplies to the Department of Defense (DoD), and they have sponsored Lincoln Laboratory to help them reduce the frequency of mishaps across their fleet. A **mishap** is an incident on board a ship usually involving damage or injury. Mishaps vary in scope, type, and severity

Project Goal: Develop accurate and interpretable models capable of predicting when mishaps occur and assisting in diagnosing their causes.

Baseline Model: MSC Currently uses a model called the Probability Risk Indicators for Safety Management (PRISM) to calculate values known as PRISM scores using five features (PRISM features) for each ship. Ships are rank-ordered and assigned risk levels on a quarterly basis.

Low Risk
Ship A, Ship B, Ship C, …

Medium Risk
Ship D, Ship E, Ship F, …

High Risk
Ship G, Ship H, Ship I, …

Data

- Log of mishaps from July 2015 ~ August 2020 (binarized later)
- Training (2015~2018), Testing (2019)
- PRISM Features
  - Days of overdue relief (days beyond initial assignment)
  - Critical billets' time on board (key leaders' time on ship)
  - Percentage of excessive overtime hours
  - Percentage of maintenance incomplete for ship
  - Number of mishaps in a previous period

Month	Mishap	Ship	Ship Class	Days of Overdue Relief	CBTOB	Excess Overtime	Maint. Incomplete	Previous Mishaps
MM/YYYY	0/1	Ship X	Class Y	#	#	%	%	#
⋮								

Modeling

1. Transform PRISM into a binary classification model

Compute	Compute PRISM scores at monthly level
Divide	Divide every score by the maximum score value
Compare	Generate ROC curves and compare models

2. Resample training data for balanced class distribution

Duplicate entries with more than one mishap by number of mishaps after binarization

Month	Ship	Mishap	…
MM/YYYY	Ship X	3	…

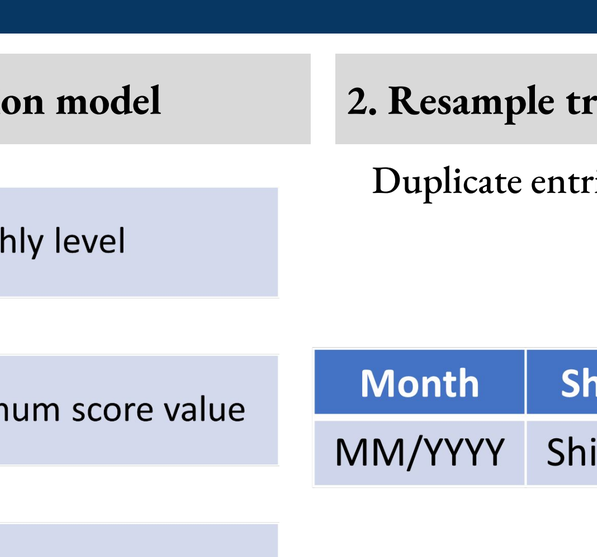
→

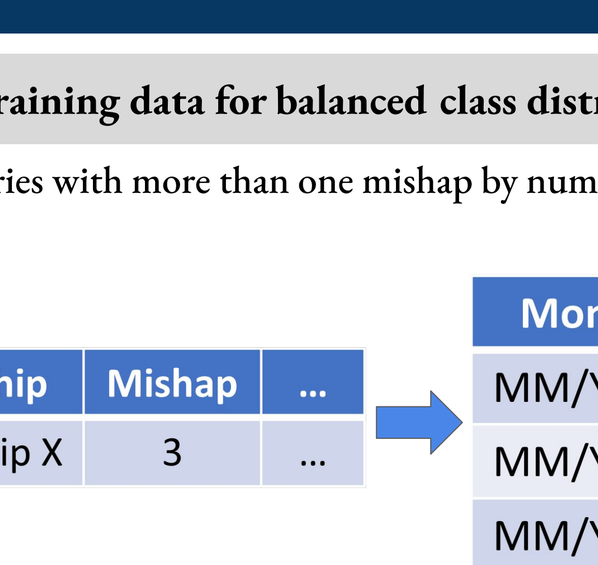
Month	Ship	Mishap	…
MM/YYYY	Ship X	1	…
MM/YYYY	Ship X	1	…
MM/YYYY	Ship X	1	…

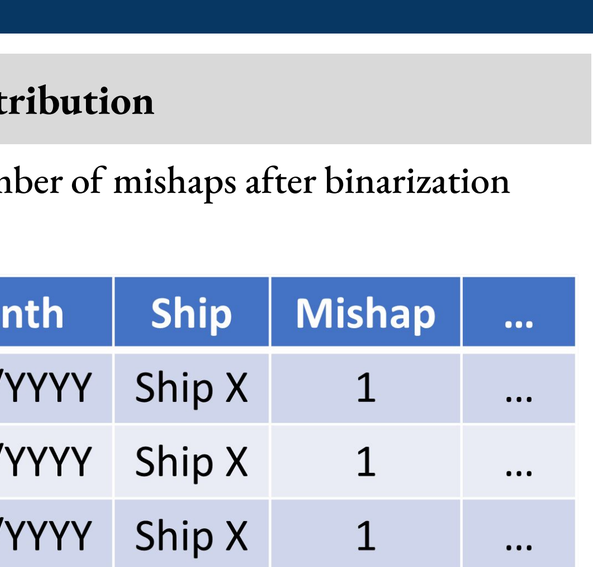
3. Build several predictive models, evaluate performance & feature importance

Models were built at the ship class level. Ship classes categorize ships by form and function similar to automobile types. Our best models in terms of AUC were logistic regression and CART. Logistic regression was a higher performing model, but CART offered interpretable trees. Both models identified the number of mishaps in a previous period (90 days and 30 days) to be the most significant features in predicting future mishaps.

	Ship Class A	Ship Class B	Ship Class C
PRISM	0.602	0.527	0.621
Logistic Regression	0.835	0.831	0.814
CART	0.712	0.791	0.711
OCT	0.700	0.715	0.519
Random Forest	0.738	0.718	0.586







Potential Business Impact and Conclusion

Our best models outperformed PRISM by a considerable margin. We estimate they can **predict 82% of mishaps** before they occur. Our models also highlighted the importance of specific features and transformed them in a way that led to improved machine learning models. MSC can use this information to operate their ship classes differently. They can also provide these insights to leaders on board ships so they can take measures to reduce mishap occurrence.

Mishaps Predicted:
~82%

This material is based upon work supported by the United States Transportation Command under Air Force Contract No. FA8702-15-D-0001. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the United States Transportation Command.