NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

**PREDICTING FUTURE DESTINATIONS
OF TACTICAL UNITS**

by

Jun H. Kim

June 2023

Thesis Advisor:	Robert A. Koyak
Second Reader:	Robert L. Bassett

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704-0188*

1. AGENCY USE ONLY ***(Leave blank)***	**2. REPORT DATE** June 2023	**3. REPORT TYPE AND DATES COVERED** Master’s thesis

4. TITLE AND SUBTITLE PREDICTING FUTURE DESTINATIONS OF TACTICAL UNITS	**5. FUNDING NUMBERS**
6. AUTHOR(S) Jun H. Kim	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A	**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release. Distribution is unlimited.	**12b. DISTRIBUTION CODE** A

13. ABSTRACT (maximum 200 words)

The purpose of this thesis is to apply machine learning techniques towards predicting the future destinations of tactical units that move in a known road network. These units are modeled after standard field artillery batteries. Each battery is made up of eleven vehicles: four launcher vehicles, four reloading vehicles, two support vehicles, and one command control vehicle. Data was generated by the Modeling Virtual Environments and Simulation (MOVES) institute at NPS.

There are two study questions: Can machine learning models accurately predict the future destinations of tactical vehicles? What is an adequate level of prediction accuracy for use in tactical applications?

Of the current machine learning techniques, we use random forests and neural networks for destination prediction. Overall, our random forest achieves 38.9 percent prediction accuracy while our neural network achieves 43.2 percent prediction accuracy.

There are four immediate directions for future research following this thesis. They are further investigation of prediction modeling, using data with measurement error collected on irregular time intervals, modeling with real world data, and multi-domain modeling.

14. SUBJECT TERMS Modeling Virtual Environments and Simulation, MOVES, random forest, neural networks, machine learning, supervised learning, tactical vehicle, battery	**15. NUMBER OF PAGES** 53
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

PREDICTING FUTURE DESTINATIONS OF TACTICAL UNITS

Jun H. Kim
Major, United States Army
BS, Texas A&M University, College Station, 2010
MBA, University of Maryland, 2019

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
June 2023

Approved by: Robert A. Koyak
Advisor

Robert L. Bassett
Second Reader

W. Matthew Carlyle
Chair, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

The purpose of this thesis is to apply machine learning techniques towards predicting the future destinations of tactical units that move in a known road network. These units are modeled after standard field artillery batteries. Each battery is made up of eleven vehicles: four launcher vehicles, four reloading vehicles, two support vehicles, and one command control vehicle. Data was generated by the Modeling Virtual Environments and Simulation (MOVES) institute at NPS.

There are two study questions: Can machine learning models accurately predict the future destinations of tactical vehicles? What is an adequate level of prediction accuracy for use in tactical applications?

Of the current machine learning techniques, we use random forests and neural networks for destination prediction. Overall, our random forest achieves 38.9 percent prediction accuracy while our neural network achieves 43.2 percent prediction accuracy.

There are four immediate directions for future research following this thesis. They are further investigation of prediction modeling, using data with measurement error collected on irregular time intervals, modeling with real world data, and multi-domain modeling.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

ADS-B	Automatic Dependent Surveillance-Broadcast
AIS	Automatic Identification System
BFT	Blue Force Tracker
C2	Command and Control
CNN	Convolutional Neural Network
EDA	Exploratory Data Analysis
ISR	Intelligence, Surveillance, Reconnaissance
JCR	Joint Capabilities Release
JRTC	Joint Readiness Training Center
LSTM	Long Short-Term Memory
MDCOA	Most Dangerous Course of Action
MLCOA	Most Likely Course of Action
MOVES	Modeling Virtual Environments and Simulation
NPS	Naval Postgraduate School
NTC	National Training Center
POTUS	President of The United States
RNN	Recurrent Neural Network

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

The goal of this thesis is to provide destination prediction of tactical vehicles in a known, closed road network. These tactical vehicles are modeled after wheeled, field artillery units. The Modeling Virtual Environments and Simulation (MOVES) institute at the Naval Postgraduate School (NPS) modeled and generated the data within a fictional scenario. The scenario features various locations and events that a typical field artillery unit would encounter in a deployed environment. The military organization consists of a single regiment made up of two battalions, each with four batteries, and each battery containing eleven vehicles. Within each battery, there are four launcher vehicles, four reloading vehicles, two support vehicles, and one command and control (C2) vehicle. Data generation is recorded at the regiment, battalion, battery, and vehicle level. This study is centered around battery movement patterns. Each scenario is broken down into smaller trips, where there is only one prior location and future destination. The predictor variables in the model fitting describe various positional attributes of each battery. The response variable is the destination location of each trip.

This thesis examines two primary study questions.

1. Can machine learning models accurately predict the future destinations of tactical vehicles?
2. What is an adequate level of prediction accuracy for use in tactical applications?

This thesis is constrained to using only generated data from the MOVES institute. As a result, some limitations are imposed. The first limitation is that the data lack measurement or sensor error. Perfect data collection in the real operating environment is not realistic. The second limitation is that the generated data is sufficient in size to fit our models. In new combat scenarios, the data can be sparse or not available.

To predict the future destinations of these tactical units, we use two machine learning, supervised techniques: random forests and neural networks. To objectively compare these two models, we derive two criteria to determine success in destination prediction. Each model fits a probability to each location within each minute interval in a trip. The first criterion is if over half of the correct location assignments have above 80 percent probability. The second criterion is whether the model assigns over 80 percent probability to the correct destination in the last three minutes of a trip. A model must meet both criteria to qualify as successful. The random forest achieves 38.9 percent successes of all trips in the validation set, while the neural network achieves 43.2 percent successes. A real-world scenario is considered using these two criteria. Each trip is reduced to only the first five minutes. In real situations, decision makers must decide on actions before the enemy completes their movements. In this scenario, the decision maker has a five-minute window before deciding. Random forest performs at 19.1 percent prediction accuracy while the neural network achieves 33.9 percent prediction accuracy. This is the upper bound of prediction accuracy for the time constrained scenarios. With the introduction of error and noise, prediction accuracy is likely to decrease.

This thesis sets the upper bound of destination prediction by using perfect data. The following are areas for future research based on our thesis: further investigation of prediction modeling, handling data with measurement error collected on irregular time intervals, modeling with real world data, and multi-domain modeling. The first area is to improve prediction accuracy by further prediction modeling. The second area is to introduce error terms that resemble the difficulties of real-life data collection and aggregation. Battlefield sensors are not perfect and have limitations. The third area is to model with real world data from actual combat deployments and training. The last area can be a replication of our study to the other war fighting domains: sea, sub surface, and air. Motion profiles and movement behavior are equally relevant in each of these domains. Tactical level models can inform strategic level decision making.

ACKNOWLEDGMENTS

I would like to give all the glory to God in that all I do. Next, I am eternally grateful to my thesis advising team, Dr. Koyak and Dr. Bassett. They have given invaluable guidance, patience, and mentorship in this difficult journey. It has been a true blessing to learn from such world-class professors. Lastly, I would like to thank my family for their constant support during my studies here at NPS.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

The land domain is a critical aspect of the joint battlespace. With advanced intelligence, surveillance, reconnaissance (ISR) and over the horizon communications, commanders can achieve battlefield effects from beyond the area of operation. However, the tactical level fight still requires ground command and control (C2). Central to the ground tactical fight is indirect fires. Close air support and other long range precision fires are costly and are considered precious resources. Army and Marine Corps field artillery capabilities are tightly integrated with modern day infantry combat units. Indirect firing units bear tactical considerations to both friendly and enemy forces. This makes destination prediction important for several reasons. Prediction based on past behaviors can aid decisions about deployment of combat power on the battlefield. However, enemy activity is difficult to profile and forecast. Even in a known road network, vehicles can travel off road and exhibit erratic behavior. Based on motion profiles, our work aims to lessen the fog of war through future destination prediction.

A. THESIS PURPOSE

The purpose of this thesis is to use machine learning techniques towards predicting the future destinations of tactical units. The dataset that we use was generated by the Modeling Virtual Environments and Simulation (MOVES) institute at the Naval Postgraduate School (NPS). In these scenarios, tactical units travel on a known road network, to various destinations, based on scenario transition probabilities. Predicting the destination of a unit is our main area of focus. Accurate prediction of unit destination provides valuable information to tactical decision makers. They must precisely maneuver their formations to anticipate and defeat enemy forces. Destination prediction helps them distinguish between threatening and non-threatening activity. Accurate prediction is also important to achieve optimal placement and alignment of sensors throughout the battlefield.

B. STUDY QUESTIONS

This thesis examines the various attributes of past and present motion profiles on a road network. Attributes include both spatial and time components. There are two study questions that we address in this thesis.

1. Can machine learning models accurately predict the future destinations of tactical vehicles?
2. What is an adequate level of prediction accuracy for use in tactical applications?

C. SCOPE, LIMITATIONS, AND ASSUMPTIONS

Our main research interest is to determine the future destination of tactical units that are traveling within a closed, known road network. These tactical units are called batteries. Each battery has eleven wheeled vehicles: four launchers, four reloaders, two support vehicles, and one C2 vehicle. The generated scenario contains a single regiment that is comprised of two battalions, where each battalion contains four batteries. Each battery moves independently and is the smallest grouping of vehicles in this scenario.

The dataset has some limitations. The data generation did not include any type of error. This means that each set of attributes was perfectly recorded and regarded as the ground truth. Under battlefield conditions, sensors are prone to physical conditions that would result in imperfect data. Sensors would not record in perfect one-minute increments, nor can they record on every vehicle on the road network. To overcome these limitations, this thesis provides analysis on the best-case scenario of battery motion. Under ideal datasets, our analysis is the upper limit of destination prediction. When real world data is applied, these models should be less accurate relative to perfect data conditions.

D. THESIS ORGANIZATION

The remainder of the thesis is contained within four chapters. Chapter II, Background, covers a literature review of previous works from the civilian and military sectors. Chapter III, Data Preparation and Methodology, explores the techniques and technical background used in this thesis. Chapter IV, Results, provides findings and analysis from this study. Chapter V, Conclusions, covers answers to the study questions, thesis considerations, and future work opportunities.

II. BACKGROUND

This chapter is a broad overview of civilian and military work that relates to our study. These share similarities and offer background information to our thesis. Each of these academic works feature elements of destination prediction based on spatial and time attributes. Each work uses different solutions to answer their research objectives.

A. MOTION ENTITIES

Entities in motion are the subject of many studies in both the military and civilian sectors. From animals in nature to commercial ships, there is great interest in extracting patterns, trends, and analysis of motion profiles. There are many applications of motion profiles that include: optimization of shipping routes, behavioral trend analysis, resource prioritization, resource allocation, and destination forecasting. In this thesis, we are analyzing the motion profiles of tactical, wheeled vehicles on a road network.

B. CIVILIAN RELATED WORKS

Lv et al. (2020) describe their process and methods for predicting taxi destinations within a road network. Their data was gathered from real taxis operating in Porto, Portugal. Their commercial application is to show passengers personalized ads based on their trip destination. Figure 1 is a visual graphic of this process.

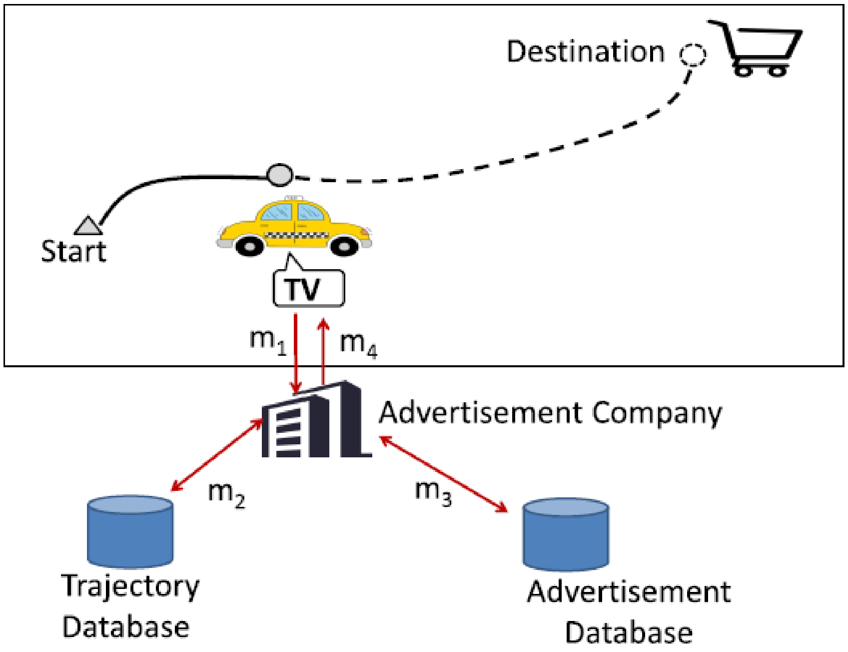
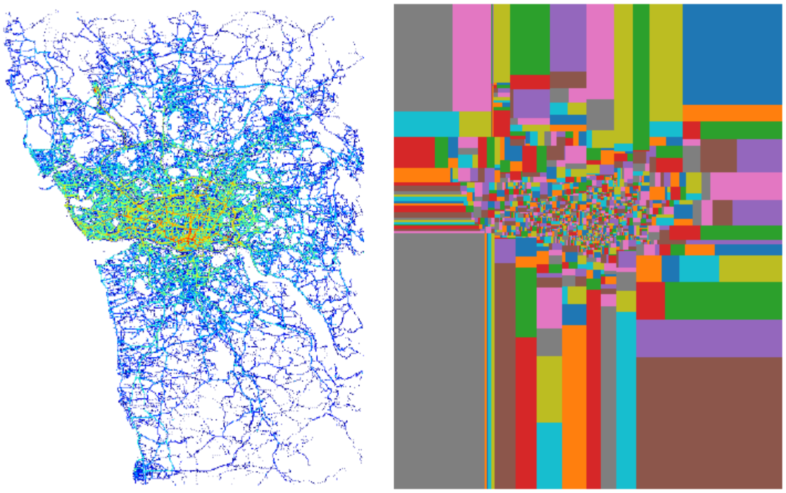


Figure 1. Ad Prediction Based on Destination. Source: Lv et al. (2020).

As the taxi travels along its path, the algorithm looks at past data to find advertisements that are associated with the destination (Lv et al. 2020). They use models with a convolutional neural network (CNN) to achieve their best accuracy in ad prediction. This commercial application has a use case for the military. Specifically, this is relevant to enemy threat analysis and situational awareness. This prediction analysis directly informs enemy items such as most likely course of action (MLCOA) and most dangerous course of action (MDCOA). Based on a suspected destination, a model can determine probable event outcomes and levels of risk.

Ebel et al. (2020) examine data that features taxi travel sequences. Their approach is to use latitude and longitude data to spatially partition the known road network. Figure 2 shows the road network made of smaller, distinct regions.



(a) Denstity map of data points (b) Discrete regions

Figure 2. Road Network Segmented into Discrete Regions. Source: Ebel et al. (2020).

Their final model combines this type of space partitioning with a neural network featuring long short-term memory (LSTM) layers. By using these partitioned regions, they examine the sequences of taxi trips. This sequence of visited regions is suited for destination prediction using the LSTM model.

Endo et al. (2017) frames their problem using trajectories, meaning each trip has a start and end location. Based on the past trajectories, they aim to train a model to predict the destination, as shown in Figure 3.

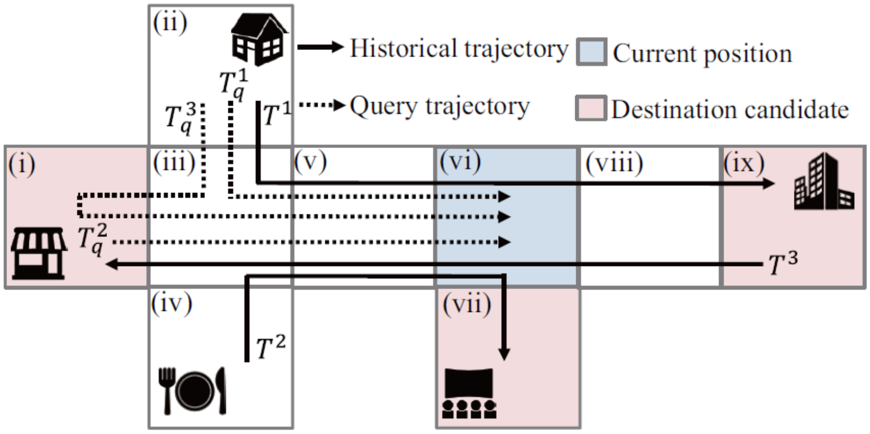


Figure 3. Trajectories for Destination Prediction. Source: Endo et al. (2017).

Using these trajectories, Endo et al. (2017) use a recurrent neural network (RNN) to predict trajectory destination. The recurrent property is utilized because the model performs better when it can train on more past trajectories. Each of these civilian studies has differing approaches to this destination prediction yet applied varying solutions.

C. MILITARY RELATED WORKS

Young (2017) focuses on the Automatic Identification System (AIS) data of shipping vessels leaving the Los Angeles port. His thesis aims to predict vessel destination and route, using various machine learning techniques. His final models use the following attributes: speed, latitude, longitude, and course direction. Figure 4 shows a visual example of vessel routes to an example prediction.

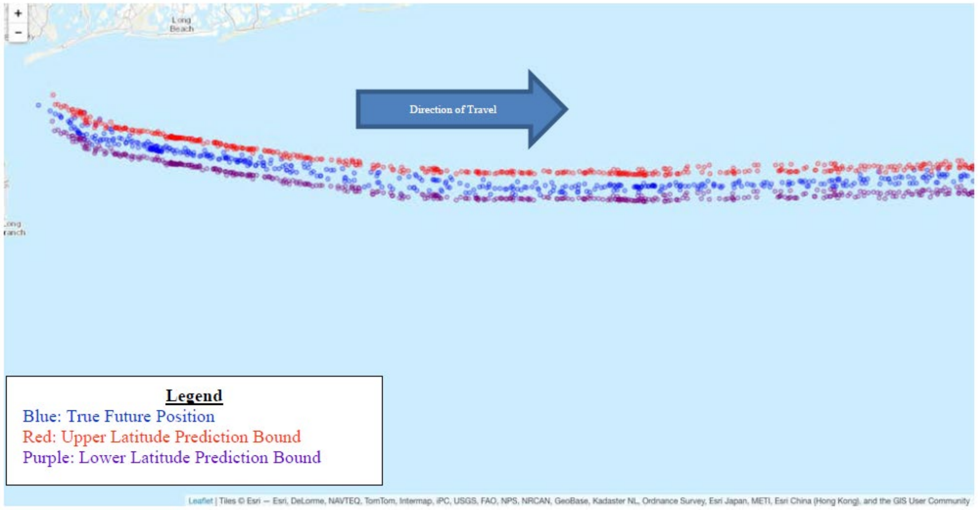


Figure 4. Upper and Lower Bound Prediction of a True Route. Source: Young (2017).

His models fit an upper and lower bound around the true route going towards the destination. In the case above, the prediction bounds contain the truth throughout the whole route.

Lee (2020) explores the use of small datasets to predict aircraft destinations. His thesis uses Automatic Dependent Surveillance-Broadcast (ADS-B) data, containing aircraft coordinates, heading, speed, ID, time, and squawk. His scope is specific to only flights from Dublin to European regions and associated with Ryanair flights. His final model features Bayesian estimation that gives a prediction based on the flight ascent, cruise, and descent. Figure 5 is a graphical depiction of Ryanair flight paths.

Figure 5. Ryanair's Flights from Dublin. Source: Casey (2020).

This thesis shares numerous similarities to our study. Each route has various spatial and time attributes. Specifically, we also utilize heading angles and time predictors to determine the destination.

Pham (2019) looks at AIS data of cargo ships that enter the Baltic Sea enroute to various locations. He uses the following predictors to fit his model: distance to each port, bearing to each port, and difference in bearing from the entry point to each port. Figure 6 shows point of entry from the southwest location and routes to destinations.

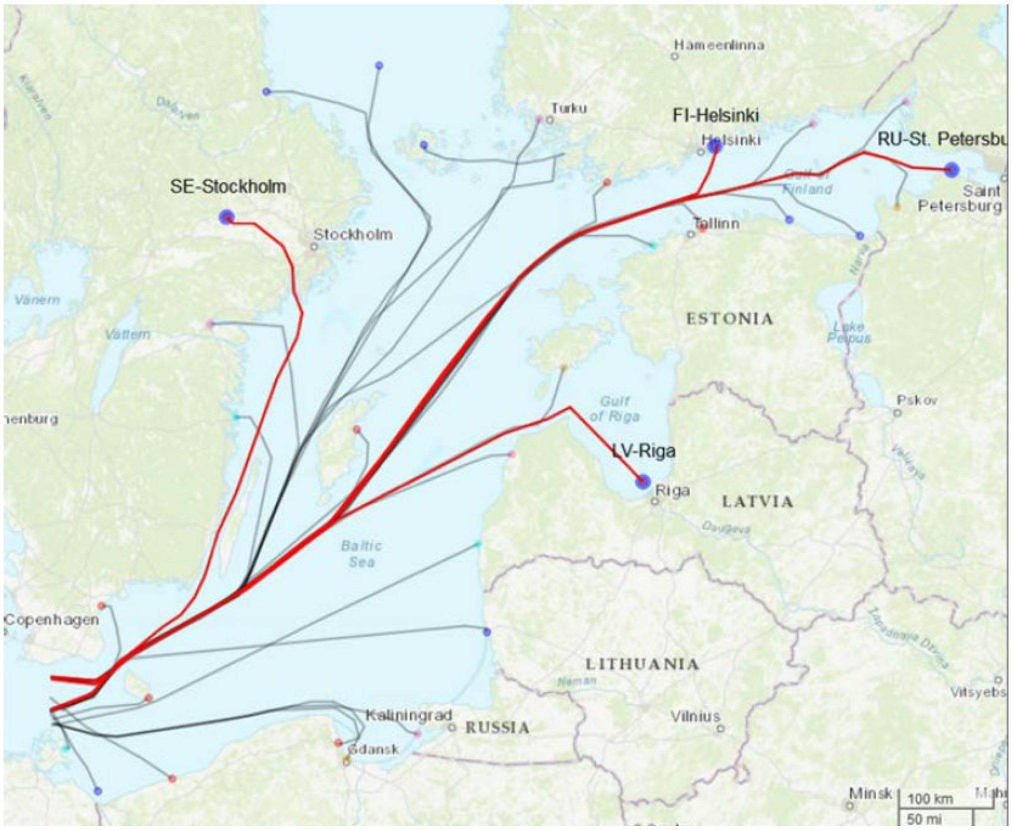


Figure 6. Possible Destinations in the Baltic Sea. Source: Pham (2019).

In his final models he uses random forest to predict the destinations of these vessels once they enter the Baltic Sea. We also use bearing angle differences in our set of predictors along with random forests.

III. DATA PREPARATION AND METHODOLOGY

In this chapter, we conduct exploratory data analysis (EDA), data preparation, and techniques for our final models. These final models answer our study questions for this thesis.

A. DATA DESCRIPTION

The dataset was generated by the MOVES institute. They generated eighty distinct scenarios for every battery of tactical vehicles, each at 3,960 minutes duration. Each scenario has over 800,000 data points for every battery. The scenario unit task organization is as follows: one single regiment contains two battalions, comprised of four batteries, and within each battery are four launcher vehicles, four reloading vehicles, two support vehicles, and one C2 vehicle (CMIS 2022). The vehicles follow a generic configuration found in typical field artillery units. In all scenarios, vehicles travel at standard road speeds. As shown in Figure 7, each battery maneuvers throughout a known road network that features training, garrison, and attack sites.

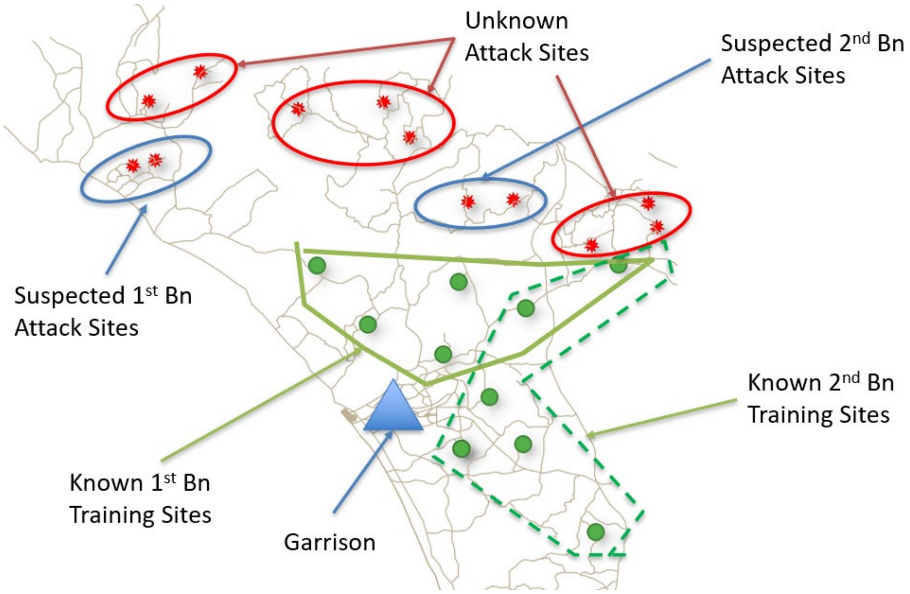


Figure 7. Map With All Destination Locations. Source: CMIS (2022).

The response variable for this thesis is the destination of a moving battery. There are 24 destinations: garrison, 10 training sites, one known attack site, 11 unknown attack sites and one destination that is unclassified. The one known attack site is a destination that is known to friendly forces as a destination to engage the enemy. The unknown attack sites are destinations that are not known to friendly forces at the start of the scenario. The unclassified destination describes locations that are neither of these other locations (e.g., stopping for fuel in the middle of a trip).

B. DATA PREPARATION

To perform analysis and prediction, significant data cleaning and conditioning is required. The R programming language (R Core Team 2023) is used to perform basic exploratory data analysis and preparation. The first step is to gather all the entity level data together into a single dataset. Using the longitude and latitude, we derive two predictors: Start Location and Destination Location. The start location in a trip adds significant information to our models. Destination location is our response variable. The next step is to separate the data into only motion segments referred to as trips. We retain only the trips for our analysis. Non-motion activity is not included in model data. Each trip is the movement of a battery travelling from one location to the next. It is a subset of the unit's overall scenario movement. The dataset is randomly split based on the total number of trips (6,853). The training set has 70 percent (4,797) of all trips. The testing and validation sets each have 15 percent (1,028) of all trips.

After splitting the data, we conduct EDA on the trip length of each data set. The training set has an average trip length of 33.85 minutes with a standard deviation of 17.29 minutes. Most of these trips are between 10 to 60 minutes in length. As shown in Figure 8, the training set trip lengths look bimodal with right skewness.

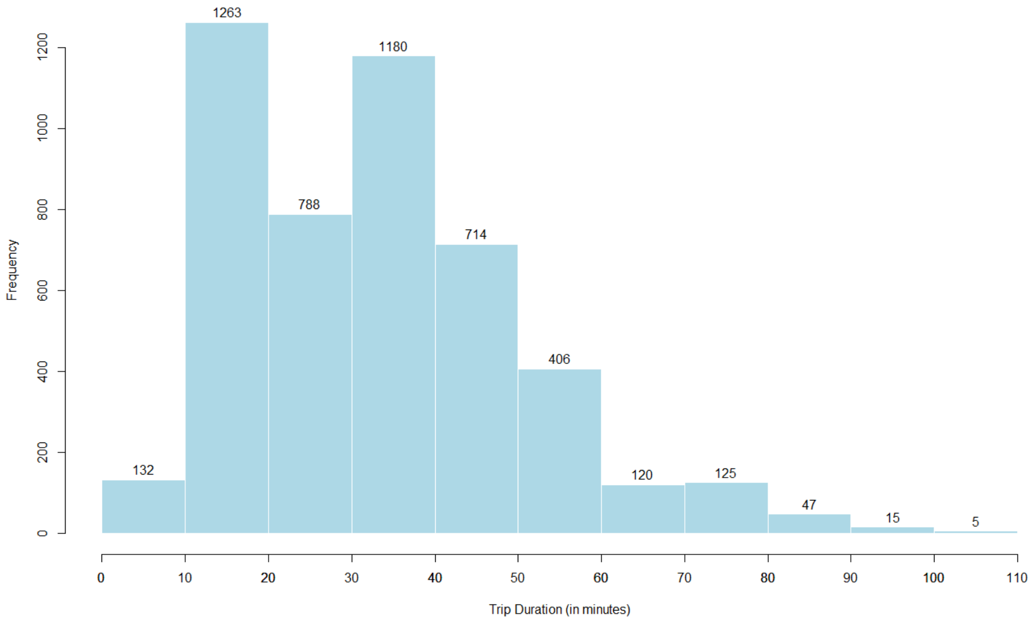


Figure 8. Training Dataset Trip Lengths

The testing dataset has 1,028 trips with an average length of 33.00 minutes with a standard deviation of 16.65 minutes. Most trips are between 10 to 50 minutes. As shown in Figure 9, this set is also bimodal with right skewness.

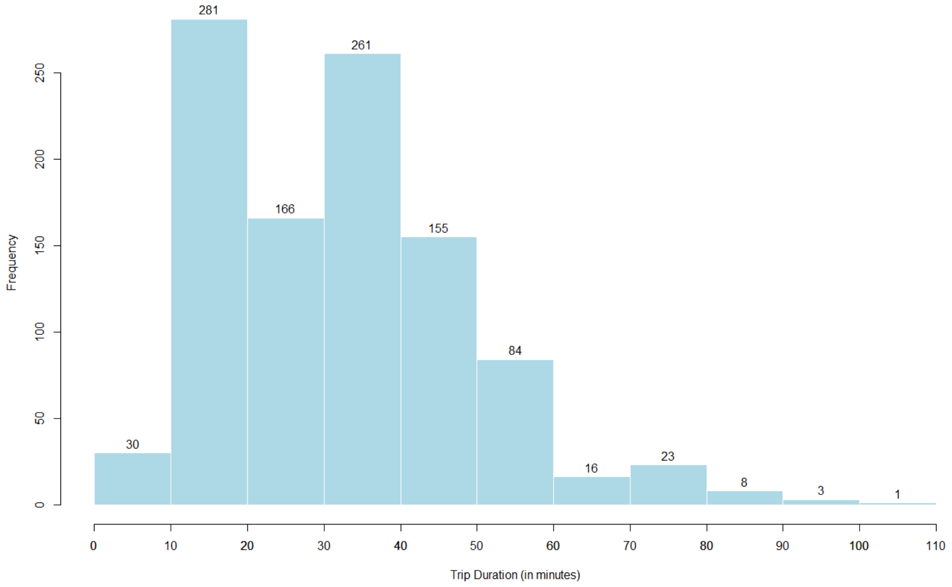


Figure 9. Testing Dataset Trip Lengths

The validation set has 1,028 trips with an average of 33.42 minutes and standard deviation of 16.76 minutes. Most trips are between 10 to 60 minutes long. As shown in Figure 10, this set is bimodal with a right skew.

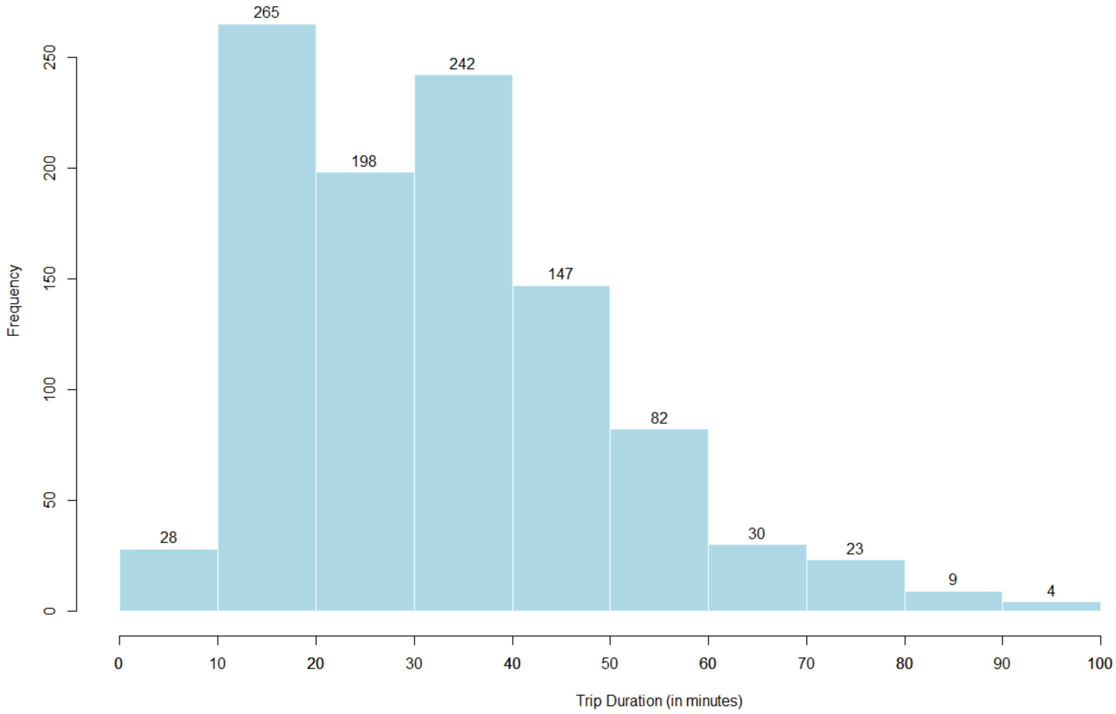


Figure 10. Validation Dataset Trip Lengths

Additional feature engineering is needed relative to battery position. Out of 24 destinations, there is one garrison, 12 attack sites, and one unclassified site. These locations are chosen because they are located at the top and bottom of the area of operation, while training sites are in the middle section. As shown in Figure 11, three distinct attributes (B,BD, D) are generated for these 14 locations.

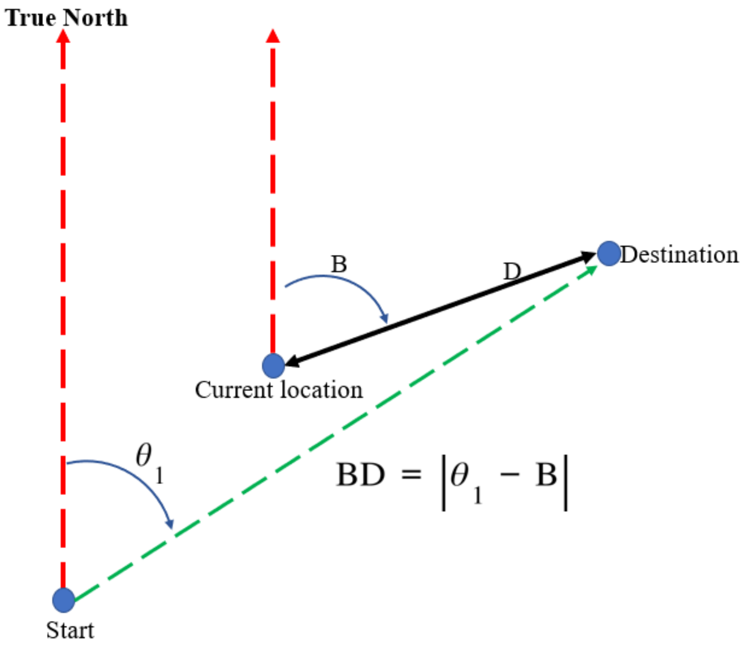


Figure 11. D, B, BD Predictors

Predictors D, B, and BD each generate a set of 14 distinct values in each increment of every trip. Attribute D is the straight-line distance from current battery location to each of the fourteen non-training destinations. Attribute B is the angle (in degrees) from true north to the angle of the destination. BD is the difference of θ_1,the angle from starting location to the destination, and the B angle. The calculation of BD is shown here:

$$\left[\left(\theta_1 - B\right) \bmod 360 + 540\right] \bmod 360 - 180$$

This calculation is the difference of θ_1 and B followed by modulus and linear operations. The modulus operation refers to the remainder of the Euclidean division of two numbers. Angles wrap around at 180 degrees (i.e., 180 degrees and -180 degrees are the same). Angles can be determined in two ways. Angles take on a negative value if measured going counterclockwise. Angles have positive values when measured in the clockwise direction. This calculation is done to correctly account for angles in our data.

C. METHODOLOGY

Because our problem is one of classification, we consider the following machine learning techniques: random forests, neural networks, convolutional neural networks , and long short-term memory layer in recurrent neural networks.

A random forest is best described as a collection of decision trees where each node is an attribute with further nodes and branches (Chollet et al. 2022). Each decision tree produces a majority vote for one of the possible classes, which in our case is a destination. Random forests offer some key benefits such as reduced risk of overfitting, flexibility between classification and regression, and ease of determining feature importance. Some drawbacks are that this algorithm takes considerable time to run and will require large memory storage when using large datasets.

Neural networks are widely used supervised machine learning techniques. Chollet et al. (2022) explains how neural networks determine how each of the predictors should be weighed, using activation and loss functions, in a layered structure. In a neural network, the output of previous layers feed into the next. The activation function is a non-linear function applied to the layers. It uses linear combinations of weights, to determine if its outputs should go into the next layer. This happens at each of the nodes in this framework. The loss function measures the distance of the weighted model prediction from the truth. Chollet et al. (2022) describes how convolutional neural networks operate in the same manner but differ by how the inputs are fed into the model. This type of network works best with image, speech, or audio signal inputs. Recurrent neural networks use sequences of data to make predictions and are best suited for natural language processing, speech recognition, and time series problems (Chollet et al. 2022).

D. MODEL PARAMETERS

We use R (R Core Team 2023) for our model fitting and code execution. For the random forest, we use the ranger package (Wright 2023) with a mandatory decision tree split on the relative time. We always want to retain this predictor in each of the decision trees in the random forest. Finally, we are looking to accurately predict the destination location, based on the past and present information.

For neural networks, we use the keras package (Chollet et al. 2022). Our final neural network requires specific parameters that are optimized for categorical prediction: activation functions, loss function, and optimization algorithms. We choose the following parameters based on similar examples by Chollet et al. (2022).

In each layer, we either use the activation function relu or softmax. These functions decide what value will go into the next layer (Chollet et al. 2022). Then we use the categorical cross entropy loss function and Adam optimization algorithm. The Adam optimization algorithm is a scheme to allow for better gradient propagation. The categorical cross entropy loss measures the distance between the predicted and actual response label.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. RESULTS AND ANALYSIS

After getting our data into usable formats, we execute our final models. There are a total of forty-eight predictors. Table 1 is the complete listing of predictors used for both random forest and neural network models.

Table 1. All Final Predictors Used in Random Forest and Neural Network Models

Name	Predictor Type	Description	Example
State	Categorical	Battery behavior	Fueling
Longitude	Numerical	East-West position of a point	20 degrees
Latitude	Numerical	North-South position of a point	40 degrees
Heading	Numerical	Direction towards destination from start of trip	50 degrees
Starting Location	Categorical	Starting location of battery	Garrison
Relative Time	Numerical	Time since start of trip (in minutes)	35 minutes
D_i	Numerical	Distance from current location to (14) non-training destinations. Each increment of every trip will have 14 predictors. ($D_1,\ldots,D_{14}$)	$D_1 = 500$ meters
B_i	Numerical	Direction from current location to (14) non-training destinations. Each increment of every trip will have 14 predictors. ($B_1,\ldots,B_{14}$)	$B_1 = 37$ degrees
BD_i	Numerical	Difference in heading and B_i. Each increment of every trip will have 14 predictors. ($BD_1,\ldots,BD_{14}$)	BD = 50 – 37 = 13 degrees

A. RANDOM FOREST MODEL SELECTION

For the random forest model, we use all 48 available predictors. The random forest model uses 2,000 bootstrap samples with Relative Time as a constant decision tree splitting parameter. This model attains a destination prediction accuracy of 57.9 percent on the testing set. Fitting a random forest model of this size requires substantial local memory and several hours.

B. NEURAL NETWORK MODEL SELECTION

For the neural network, we apply two sets of models. The first set uses all 48 available predictors. The second set uses 46 predictors, excluding State and Starting Location. All models ran for 100 epochs on a personal laptop with the following specifications: three gigahertz central processing unit, 16 gigabytes of physical memory, and no dedicated graphics card. An epoch denotes a cycle of execution in the model. A total of eight variations of neural networks are considered; however, accuracy begins to drop after five layers. Model accuracy is measured by the accuracy of the predicted destination relative to the actual destination. Model loss is measured by categorical cross entropy loss as described earlier. One dimensional CNN and LSTM models are also applied. These models are measured on the testing set. Every neural network model is feed forward except the LSTM and CNN. Table 2 shows the neural network description, accuracy, loss, and training time of the variations of neural networks.

Table 2. Neural Network Models

Description	Prediction Accuracy	Model Loss	Training Time (minutes)
46-numeric predictors, 3-layer NN	57.5%	1.190	9
46-numeric predictors, 4-layer NN	57.7%	1.248	11
46-numeric predictors, 5-layer NN	57.5%	1.406	19
46-numeric predictors, 2-layer LSTM	58.2%	1.175	17
46-numeric predictors, CNN	58.6%	1.257	86
46-numeric predictors, 2 categorical 3-layer NN	59.9%	1.438	9
46-numeric predictors, 2-categorical predictors, 4-layer NN	59.3%	1.344	10
46-numeric predictors, 2-categorical predictors, 5-layer NN	58.4%	1.191	15

C. RESULTS

Table 2 shows that the neural network with the highest prediction accuracy is the model with 46 numeric and two categorical predictors, employing three layers with 24 units. It attains classification accuracy of 59.9 percent.

Once the best neural network is found, the two models (random forest and neural network) need to be applied to destination prediction. Depending on the twenty-four destinations, the models perform differently. In this section, we use the validation dataset (1028 trips) to generate our predictions and results. Figure 12 shows the difference between the two models on a trip that is 50 minutes long.

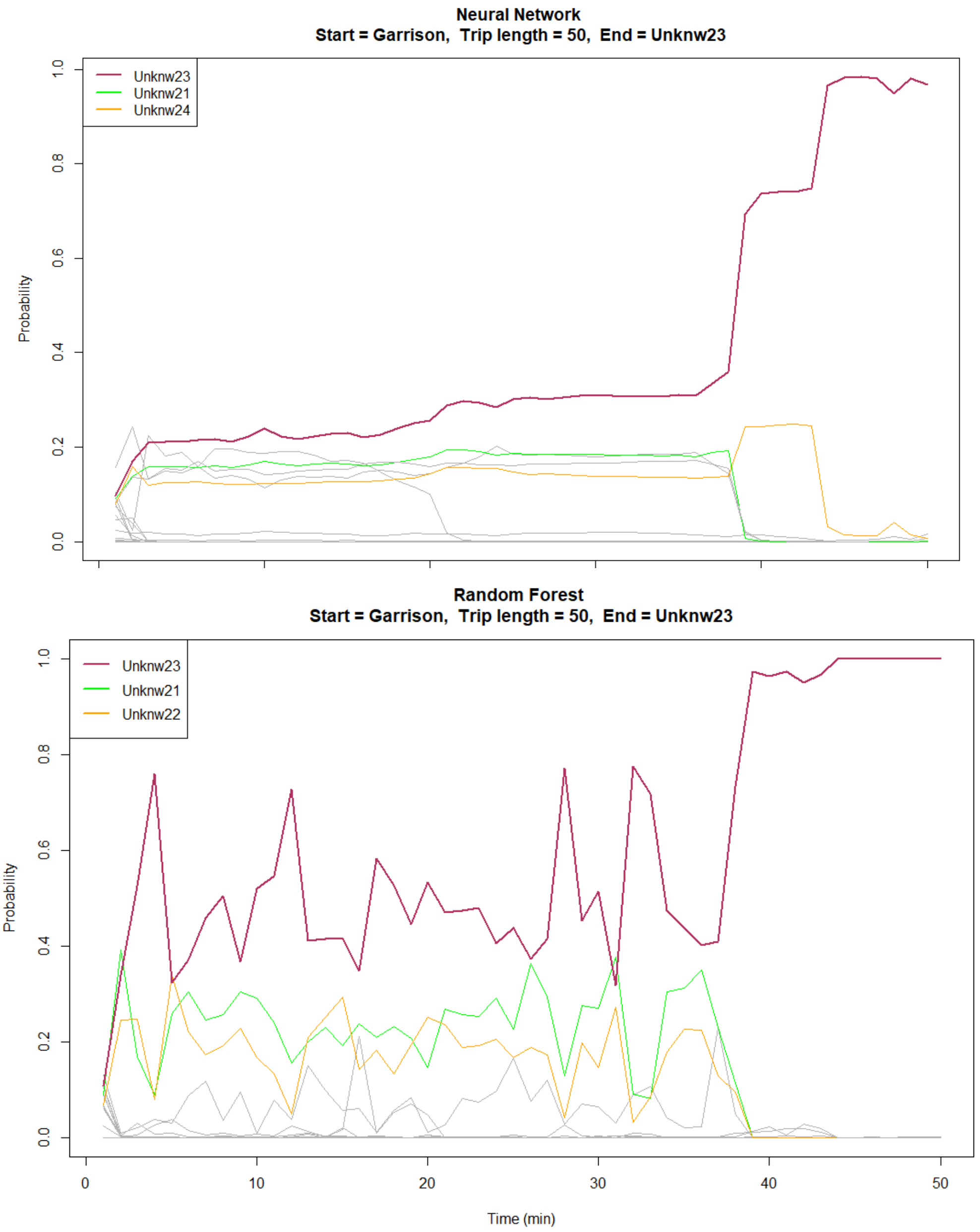


Figure 12. Attack Site Destination

In Figure 12, each model is assigning probabilities to destinations on each minute interval. As the battery moves from garrison to site Unknw23, the predictions are represented by the assigned probabilities of each site.

From time 0 to 40 minutes, both models have yet to reach 80 percent probability. Both models eventually converge on a probability of above 80 percent towards the end of the trip. Figure 13 shows another trip that is headed towards a training site.

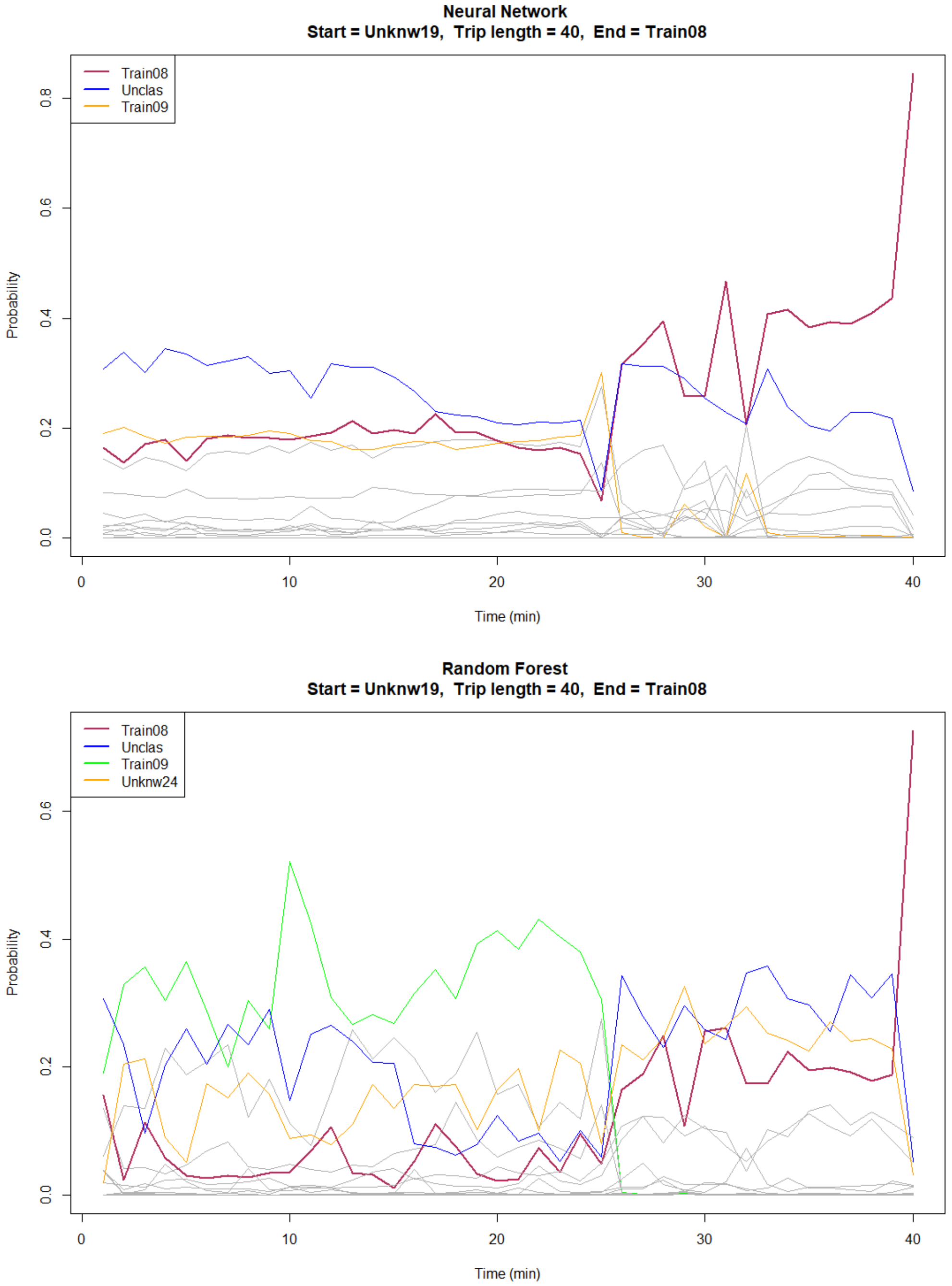


Figure 13. Training Site Destination

In Figure 13, neither model converges definitively on a destination through the first 25 minutes of the trip. After that, the neural network model probability is greater for the correct destination. Neither model, however, exceeds 80 percent probability.

Figure 14 shows the relative performance of each model with all trips in the validation set that last at least 33 minutes, which encompasses a total of 462 trips.

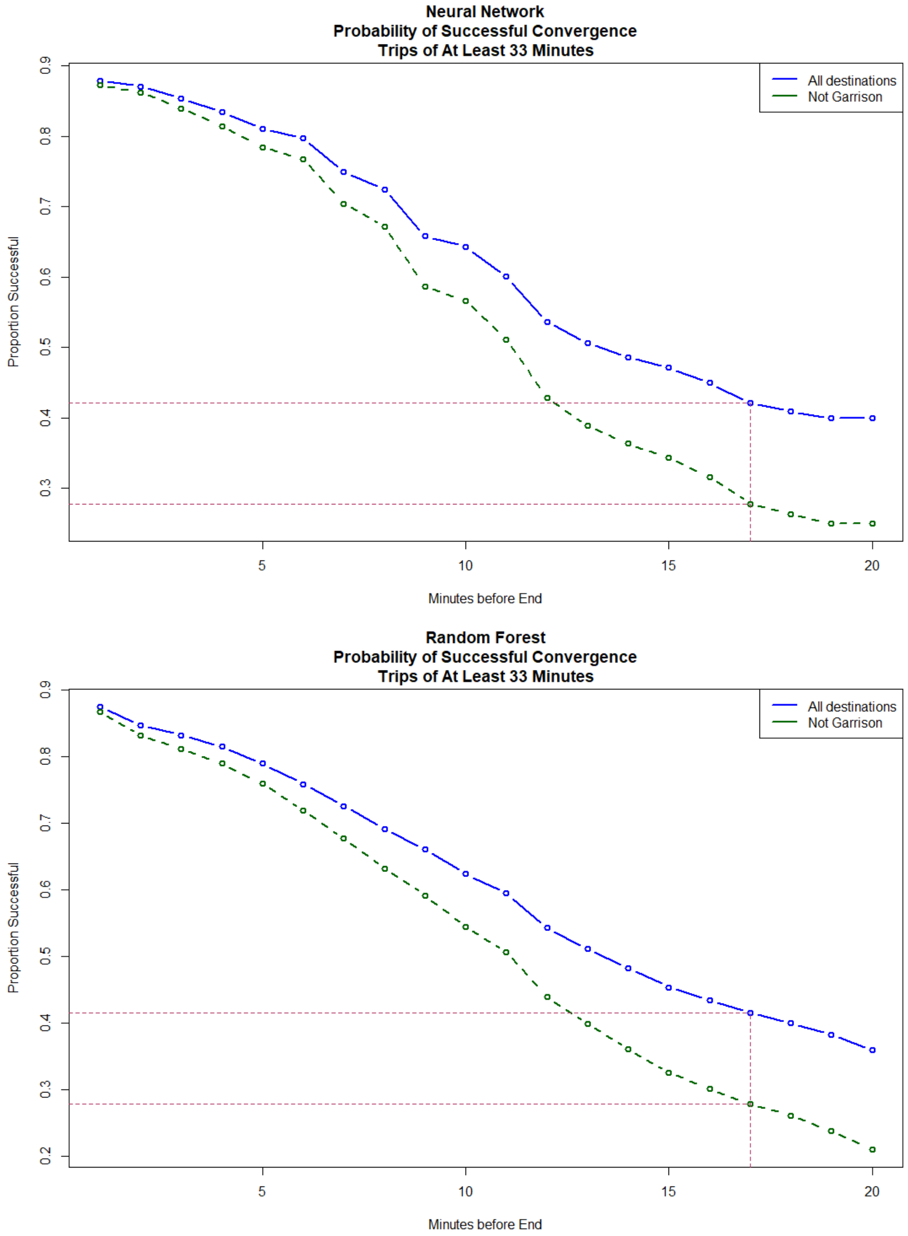


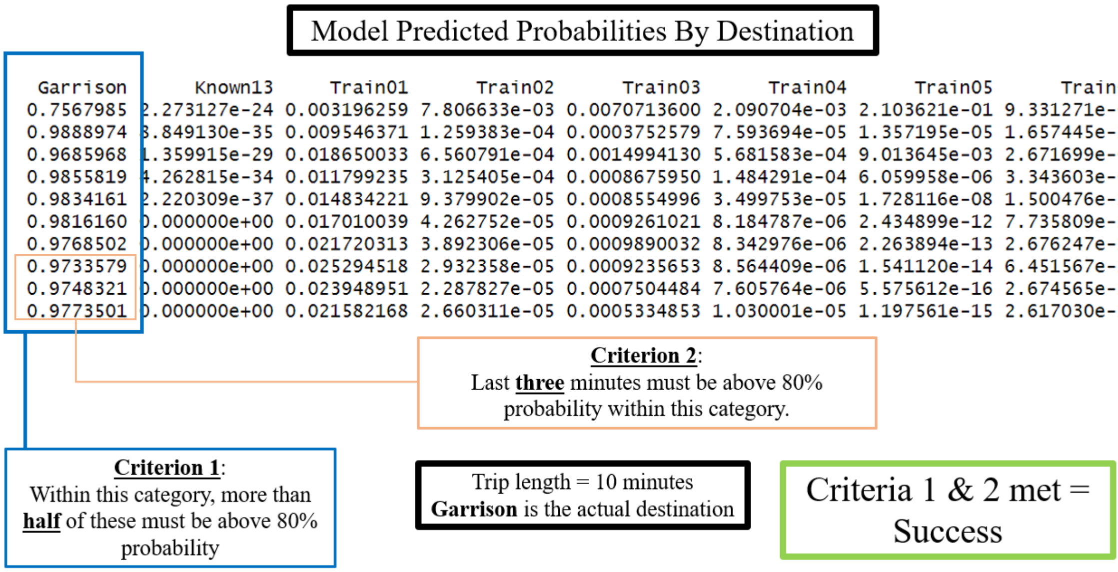
Figure 14. Performance of Models for Trips at Least 33 Minutes

Both models perform similarly throughout the 33-minute window. The horizontal axis is the time remaining before the destination is reached. At 17 minutes prior to the end of a trip, both models are at about 40 percent accuracy for all destinations, but just under 30 percent for non-garrison destinations. The farther the battery is from the destination, the lower the probability of successful prediction. The next section examines predictions across all the trips in the validation dataset (1028 trips).

D. ANALYSIS

To apply these models to every trip, we need appropriate metrics to assess overall prediction accuracy. We derive two criteria to define the success or failure of model prediction. Each trip is broken into minute intervals. The models assign probabilities to the various destinations at each time that a trip is underway. For each trip, the first criterion is the number of times a model assigns a probability over 80 percent to the correct location. The second criterion indicates whether the correct location has the highest probability during the last three minutes of a trip. A trip is designated as successfully classified if it meets these two criteria. If either criterion is not met, the trip is designated as not successfully classified. The first criterion is based on a tactical level threshold for positive identification for direct fire engagement. This threshold is for immediate threats and split-second decision making. The second criterion is meant to confirm that the model eventually converges to the true destination.

Figure 15 shows an example of a successfully classified trip of ten-minute duration using the neural network model.



Model Predicted Probabilities By Destination

```
 Garrison      Known13     Train01      Train02      Train03      Train04      Train05      Train06
0.7567985 2.273127e-24 0.003196259 7.806633e-03 0.0070713600 2.090704e-03 2.103621e-01 9.331271e-04
0.9888974 3.849130e-35 0.009546371 1.259383e-04 0.0003752579 7.593694e-05 1.357195e-05 1.657445e-10
0.9685968 1.359915e-29 0.018650033 6.560791e-04 0.0014994130 5.681583e-04 9.013645e-03 2.671699e-08
0.9855819 4.262815e-34 0.011799235 3.125405e-04 0.0008675950 1.484291e-04 6.059958e-06 3.343603e-09
0.9834161 2.220309e-37 0.014834221 9.379902e-05 0.0008554996 3.499753e-05 1.728116e-08 1.500476e-10
0.9816160 0.000000e+00 0.017010039 4.262752e-05 0.0009261021 8.184787e-06 2.434899e-12 7.735809e-11
0.9768502 0.000000e+00 0.021720313 3.892306e-05 0.0009890032 8.342976e-06 2.263894e-13 2.676247e-11
0.9733579 0.000000e+00 0.025294518 2.932358e-05 0.0009235653 8.564409e-06 1.541120e-14 6.451567e-12
0.9748321 0.000000e+00 0.023948951 2.287827e-05 0.0007504484 7.605764e-06 5.575612e-16 2.674565e-12
0.9773501 0.000000e+00 0.021582168 2.660311e-05 0.0005334853 1.030001e-05 1.197561e-15 2.617030e-12
```

Criterion 2:
Last **three** minutes must be above 80%
probability within this category.

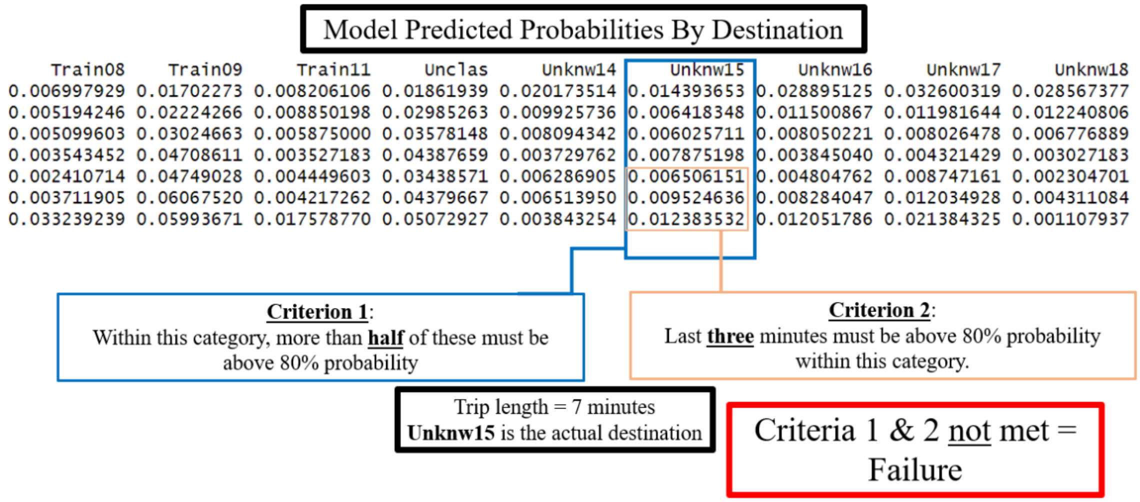
Criterion 1:
Within this category, more than
half of these must be above 80%
probability

Trip length = 10 minutes
Garrison is the actual destination

Criteria 1 & 2 met =
Success

Figure 15. Example of Successful Destination Prediction by a Neural Network Model Using Criteria 1 and 2

In Figure 15, there are 10 predictions assigned by the model with 9 of them above 80 percent, thus meeting criterion one. The last three predictions are above 80 percent, meeting criterion two. Both criteria are met resulting in successful classification. This would count as a success for the neural network. This process is also applied to probabilities attained from the random forest model. Figure 16 shows an example of a failure during a seven-minute trip.


Model Predicted Probabilities By Destination

```
    Train08     Train09     Train11      Unclas     Unknw14     Unknw15     Unknw16     Unknw17     Unknw18
0.006997929 0.01702273 0.008206106 0.01861939 0.020173514 0.014393653 0.028895125 0.032600319 0.028567377
0.005194246 0.02224266 0.008850198 0.02985263 0.009925736 0.006418348 0.011500867 0.011981644 0.012240806
0.005099603 0.03024663 0.005875000 0.03578148 0.008094342 0.006025711 0.008050221 0.008026478 0.006776889
0.003543452 0.04708611 0.003527183 0.04387659 0.003729762 0.007875198 0.003845040 0.004321429 0.003027183
0.002410714 0.04749028 0.004449603 0.03438571 0.006286905 0.006506151 0.004804762 0.008747161 0.002304701
0.003711905 0.06067520 0.004217262 0.04379667 0.006513950 0.009524636 0.008284047 0.012034928 0.004311084
0.033239239 0.05993671 0.017578770 0.05072927 0.003843254 0.012383532 0.012051786 0.021384325 0.001107937
```

Criterion 1:
Within this category, more than **half** of these must be above 80% probability

Criterion 2:
Last **three** minutes must be above 80% probability within this category.

Trip length = 7 minutes
Unknw15 is the actual destination

Criteria 1 & 2 not met = Failure

Figure 16. Example of Failed Destination Prediction by a Random Forest Model Using Criteria 1 and 2

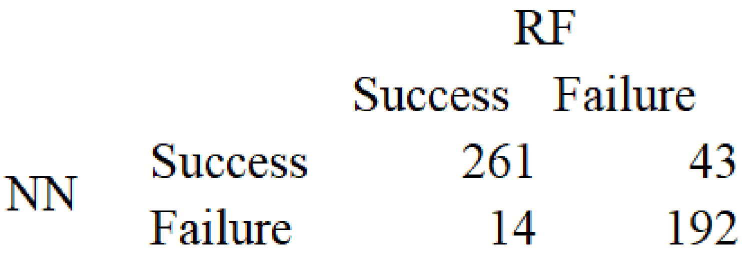
Here, the random forest model assigns the probabilities highlighted in blue for the correct destination, of these none of the 7 are above 80 percent. All of the last three predictions fail to achieve 80 percent. Since both criteria are not met, this is counts as a failure of prediction for random forest.

In the validation dataset, we classify success and failure for all 1,028 trips. Using these two criteria, we run McNemar's test (Sprent and Smeeton 2007) to determine if the success proportions from our two models are related. The null hypothesis is that the success proportions are the same for random forest and neural network models. The alternative is that they are not. The random forest model has 400 successes while neural network model had 444 successes. As shown in Figure 17, we compare the successes of each model using all 1,028 trips in the validation set.

		RF	
		Success	Failure
NN	Success	380	64
	Failure	20	564

Figure 17. McNemar's Test for Success and Failure for 1,028 Trips

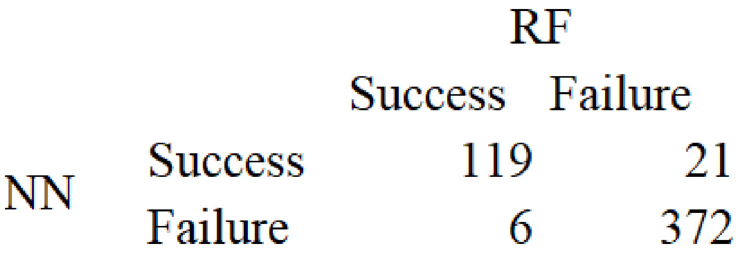
With a p-value of near zero (1.58e-06), we can reject the null at the 5 percent significance level. The test statistic is 23.048, which is compared to a chi-squared distribution with one degree of freedom. We also examine trips less than 33 minutes and apply the same test. We find the random forest has 275 successes and the neural network has 304 successes as shown in Figure 18.



		RF	
		Success	Failure
NN	Success	261	43
	Failure	14	192

Figure 18. McNemar’s Test for Trips Less Than 33 Minutes

With a p-value of near zero (1.225e-04), we can reject the null at the 5 percent significance level. The test statistic in this case is 14.754, again compared to a chi-squared distribution with one degree of freedom. This process is repeated for trips at or longer than 33 minutes. As shown in Figure 19, the random forest has 125 successes, and the neural network has 140 successes.



		RF	
		Success	Failure
NN	Success	119	21
	Failure	6	372

Figure 19. McNemar’s Test for 33 Minutes or Longer

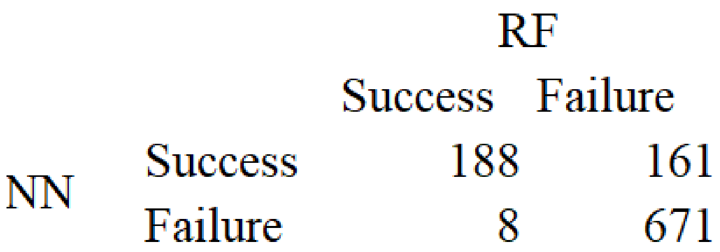
In this case, the p-value was still near zero (3.892e-03) but was higher compared to the previous tests. The test statistic is 8.333 with one degree of freedom. We can reject the null hypothesis in all three cases in favor of the alternative that the model success proportions are not the same.

For practical application, we model a scenario that is closer to real world conditions. In such operations, tactical commanders need to act quickly based on limited information. With indirect fire batteries, decisions to attack, interdict, and maneuver would require advance notice and preparation. Only minutes of observation and data collection might be available before a decision must be made. This scenario is not only applicable to ground artillery units. For example, adjacent nations in Southeast Asia maneuver their ships between training and attack postures. Some of these nations are neither allies nor coalition partners. Naval commanders must decide to interdict or change posture based on limited information. The same principles apply to certain parts of Europe at the present time. Militaries in this part of the world train and fight near their home bases. This is similar to the battery motion profiles from our analyzed scenarios.

Our hypothetical scenario uses all the validation set trips (1,028) to tabulate the total successes and failures for the random forest and the neural network. We restrict the data window to only the first five minutes of every trip since all trips in this set range in duration from 6 to 93 minutes. This scenario postulates that a tactical commander will gather five minutes of data before deciding on how to react to the situation. To consider a success or failure, we apply the same two criteria presented earlier:

- Criterion 1: Over half predictions must be over 80 percent probability on the correct destination.
- Criterion 2: Last three minutes must all be over 80 percent probability on the correct destination.

The random forest achieved 196 successes while the neural network had 349 successes as shown in Figure 20.



		RF	
		Success	Failure
NN	Success	188	161
	Failure	8	671

Figure 20. McNemar’s Test for Successes in Time Restricted Scenario

With a p-value of nearly zero (2.2e-16), we can reject the null in favor of the alternative that the success proportions are not the same. The test statistic was 138.51 with one degree of freedom.

In the non-restricted scenario, the random forest model predicts 38.9 percent successes, and the neural network model predicts 43.2 percent successes. In the five-minute restricted scenario, the random forest model predicts 19.1 percent successes, and the neural network model predicts 33.9 percent successes. For both scenarios, the neural network model performs noticeably better than the random forest model. The situation risk level and consequences would determine acceptable level of success in practice.

V. CONCLUSIONS

In the last chapter, we review our study questions along with our study limitations. Lastly, we conclude the chapter with potential areas that could prove useful to the Department of Defense.

A. REVIEW OF STUDY QUESTIONS

There are two study questions for this thesis. We answer these questions with our application of machine learning techniques and analysis.

- Can machine learning models accurately predict the future destinations of tactical vehicles?

From this study, we conclude that machine learning can aid in the destination prediction of tactical vehicles. Using random forest and neural network models, we find that the models can give between 38 to 43 percent accuracy in destination prediction.

- Can machine learning models achieve an adequate level of prediction accuracy for use in tactical applications?

A standard accuracy level cannot be assigned for all situations. In our time restricted scenario, only 19 to 33 percent prediction accuracy is achieved. The decision-making echelon will determine how much risk can be accepted. Each specific operation will determine this level of decision making. For example, a counter fire mission in well designated combat zones may be cleared by the local tactical commander. A strike in a superpower territory may require the highest level of authorization (POTUS). For these scenarios, 33 percent accuracy might be acceptable in the former. Even 99 percent accuracy would not be acceptable for the latter.

B. LIMITATIONS

The first limitation of this study concerns the data. As it was generated without measurement or sensor error, it is not reflective of actual surveillance data collection. In real world data collection, there is imperfect data obtained at irregular time intervals. The second limitation is that our models are based on using complete datasets with enough volume to satisfy the training, testing, validation set paradigm. In new or developing combat scenarios, the data may be too sparse to fit prediction models.

C. FUTURE WORK

The following are possible directions for future research based on our thesis: further investigation of prediction modeling, handling data with measurement error collected on irregular time intervals, modeling with real world data, and multi-domain modeling. Model improvement and optimization could result in higher accuracy. Combat applications would require an assessment of mission risk. Another possible direction for research would be to develop models based on actual combat situations. In the Army, there are three possible sources of information. The first two are from the two main Army training centers: the National Training Center (NTC) in Fort Irwin, CA and the Joint Training and Readiness Center (JRTC) in Fort Polk, LA. Every year, dozens of regiment size units rotate into these centers for pre-deployment training. This is the last step of pre-deployment training before these units are certified to deploy to real world combat missions. In these training rotations, units generate battery data similar to the data used in this thesis. The third possible application would be to use Blue Force Tracker (BFT) or Joint Capabilities Release (JCR) systems data from combat scenarios. These are GPS tracking systems put in most land vehicles in both training and combat deployments. The last potential area is the application to other warfighting domains. Our modelling can be generated with scenarios that contend in the sea, sub surface, and air domains. These four areas of applied work are the next evolution of study for our thesis.

LIST OF REFERENCES

Casey D (2018) Airline in focus: Ryanair. Accessed May 10, 2023, https://www.routesonline.com/news/29/breaking-news/281480/airline-in-focus-ryanair/.

Center for Mult-INT Studies (2022) Scenario Details. Unpublished report, Center for Mult-INT Studies, Monterey, CA.

Chollet F, Kalinowski T, Allaire J (2022) *Deep Learning with R Second Edition* (Manning Publications Co., Shelter Island, New York).

Ebel P, Göl I, Lingenfelder C, Vogelsang A (2020) Destination Prediction Based on Partial Trajectory Data. *2020 IEEE INTELLIGENT VEHICLES SYMPOSIUM (IV)* (Las Vegas, NV), https://doi.org/10.1109/IV47402.2020.9304734.

Endo Y, Nishida K, Toda H, Sawada H (2017) Predicting Destinations from Partial Trajectories Using Recurrent Neural Network. NTT Corporation, Yokosuka Kanagawa, JP.

Lee M (2020) Flight Destination Prediction Based on Bayesian Updating Using ADS-B Data. Master's thesis, Naval Postgraduate School, Monterey, CA.

Lv J, Sun Q, Li Q, Moreira-Matias L (2020) Multi-Scale and Multi-Scope Convolutional Neural Networks for Destination Prediction of Trajectories. *IEEE Transactions on Intelligent Transportation Systems 21(8),* doi: 10.1109/TITS.2019.2924903.

Pham C (2019) Predicting the Next Port Visit of a Vessel Using AIS Data. Master's thesis, Naval Postgraduate School, Monterey, CA.

R Core Team (2023) R: A language and environment for statistical computing, version 4.1.2. Accessed June 04, 2023, https://www.R-project.org/.

Sprent P, Smeeton, N (2007) *Applied Nonparametric Statistical Methods*, 4th ed. (CRC Press, Boca Rotan, FL).

Wright M (2023) ranger, version 0.15.1. Accessed May 10, 2023, https://rdocumentation.org/packages/ranger/versions/0.15.1.

Young B (2017) Predicting vessel trajectories from AIS data using R. Master's thesis, Naval Postgraduate School, Monterey, CA.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California