2022-09

DATA ANALYSIS AND PREDICTIVE MODEL GENERATION FOR DELAYS IN NAVY CONSTRUCTION PROJECTS

Rhea, Justin R.

Monterey, CA; Naval Postgraduate School

http://hdl.handle.net/10945/71089

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

DATA ANALYSIS AND PREDICTIVE MODEL GENERATION FOR DELAYS IN NAVY CONSTRUCTION PROJECTS

by

Justin R. Rhea

September 2022

Thesis Advisor: Hong Zhou
Second Reader: Robert L. Bassett

Approved for public release. Distribution is unlimited.

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE					*Form Approved OMB No. 0704-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC, 20503.					
1. AGENCY USE ONLY ***(Leave blank)***	**2. REPORT DATE** September 2022		**3. REPORT TYPE AND DATES COVERED** Master's thesis		
4. TITLE AND SUBTITLE DATA ANALYSIS AND PREDICTIVE MODEL GENERATION FOR DELAYS IN NAVY CONSTRUCTION PROJECTS					**5. FUNDING NUMBERS**
6. AUTHOR(S) Justin R. Rhea					
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)

Currently, Naval Facilities Engineering Command (NAVFAC) records all data on the process from application to awarding of Military Construction (MILCON) projects. This data is not utilized to increase poor performance and lack of timely results on the completion of MILCON projects. The poor performance leads to delays in deliveries to important facilities and delays in warship deployment and degradation of warfighting capabilities. NAVFAC currently has personnel investigating methods on improving the project timelines to minimize delays. Majority of the delays occur during the pre-award phase of the projects with the post-award phase causing additional delays. The purpose of this thesis is to analyze projects across multiple fiscal years from project initiation to contract award. To accomplish this, data was acquired from NAVFAC's eProjects database and analyzed using machine learning techniques as well as statistical analysis to determine a correlation between the possible causes and the delays that occurred to develop a predictive model for analyzing future project contract delays. This collection will potentially assist NAVFAC in focusing onto ongoing improvements. Reducing the delays in project awarding will further the process for reducing the overall time required to complete MILCON projects. This will shorten the amount of time that ships are in the shipyard further enhancing the Navy's undersea warfare capabilities with more submarines and other assets deployed.

14. SUBJECT TERMS Naval Facilities Engineering Command, NAVFAC, eProjects, machine learning, statistical analysis			**15. NUMBER OF PAGES** 69
			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

Approved for public release. Distribution is unlimited.

DATA ANALYSIS AND PREDICTIVE MODEL GENERATION FOR DELAYS IN NAVY CONSTRUCTION PROJECTS

Justin R. Rhea
Lieutenant, United States Navy
BSME, Old Dominion University, 2015

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN APPLIED MATHEMATICS

from the

NAVAL POSTGRADUATE SCHOOL
September 2022

Approved by: Hong Zhou
Advisor

Robert L. Bassett
Second Reader

Francis X. Giraldo
Chair, Department of Applied Mathematics

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Currently, Naval Facilities Engineering Command (NAVFAC) records all data on the process from application to awarding of Military Construction (MILCON) projects. This data is not utilized to increase poor performance and lack of timely results on the completion of MILCON projects. The poor performance leads to delays in deliveries to important facilities and delays in warship deployment and degradation of warfighting capabilities. NAVFAC currently has personnel investigating methods on improving the project timelines to minimize delays. Majority of the delays occur during the pre-award phase of the projects with the post-award phase causing additional delays. The purpose of this thesis is to analyze projects across multiple fiscal years from project initiation to contract award. To accomplish this, data was acquired from NAVFAC's eProjects database and analyzed using machine learning techniques as well as statistical analysis to determine a correlation between the possible causes and the delays that occurred to develop a predictive model for analyzing future project contract delays. This collection will potentially assist NAVFAC in focusing onto ongoing improvements. Reducing the delays in project awarding will further the process for reducing the overall time required to complete MILCON projects. This will shorten the amount of time that ships are in the shipyard further enhancing the Navy's undersea warfare capabilities with more submarines and other assets deployed.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

CNO	Chief of Naval Operations
CONUS	Continental United States
eCMS	Electronic Construction and Facility Support Contract Management System
EPG	Electronic Project Generator
FY	Fiscal Year
iNFADS	Internet Naval Facilities Asset Data Store
MILCON	Military Construction
NAVFAC	Naval Facilities Engineering Command
OCONUS	Outside the Continental United States
ReLU	Rectified Linear Unit
tanh	Hyperbolic Tangent
USW	Undersea Warfare

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

In the past decade, the United States' global naval supremacy has been dwindling at a rapid pace. Foreign countries, specifically China, have been increasing their naval power by rapidly increasing the size of their fleet. As stated by the Chief of Naval Operations (CNO) in 2018, the United States needs to increase its agility and capabilities to react and take actions (Department of Navy, 2018). Naval Facilities Engineering Command (NAVFAC) is pursuing improvement in all processes for completing projects. NAVFAC's framework for the improvements is laid out in their Strategic Design 2.0. The first component of the framework is to decrease the time it takes for projects to be awarded (Naval Facilities Engineering Command, 2019).

In this study, we focus on determining the major factors that cause projects to be delayed in the awarding process. This analysis is done by determining if a project is awarded within its assigned budget year. This is important as all projects not awarded within their budget year are reported to Congress, and often have lasting repercussions. Our analysis covers the fiscal years (FY) 2011–2021.

After performing statistical analysis, we will use machine learning algorithms, decision trees and neural networks, to generate predictive models for projects late to award. The purpose of these predictive models is to determine what projects are likely to be late to award. Further analysis into these projects will allow NAVFAC to improve project award processes.

We found that decision trees outperform the neural networks for predicting what projects will be delayed. We trained the models on FY11–20, using FY21 to test the accuracy of the models. Overall, by using decision trees we were able to create a model that is 95% accurate for predicting what projects were late, while the neural networks models were only 72% accurate.

References

Department of Navy (2018), A design for maintaining maritime superiority. Strategic Design ver 2.0, Washington, DC. https//www.navy.mil/ah_online/MaritimeSuperiority/

Naval Facilities Engineering Command (2019) External summary memo. *Strategic Design 2.0*. Washington, DC. https://hub.navfac.navy.mil/webcenter-/portal/Strategic_Design

ACKNOWLEDGMENTS

This thesis would not have been possible without the mentorship of Dr. Hong Zhou. Thank you!

Thank you, Dr. Robert Basset, for your assistance as second reader.

I would also like to thank LT Daniel Pulliam for his support and direction with using the eProjects database.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. BACKGROUND

Every year Naval Facilities Engineering Command (NAVFAC) performs statistical analysis to determine what percentage of projects are delayed in completion. In 2018, the Chief of Naval Operations (CNO) updated the Navy's expectations of the fleet. He stated that the United States needed the fleet to have increased agility across all areas (Department of Navy 2018). The increased expectations mean the NAVFAC has a more prominent role in improving fleet readiness. Most naval projects go through NAVFAC to some extent. Many NAVFAC projects, especially Military Construction (MILCON) project, are large scale, requiring substantial amounts of time, effort, and personnel to complete. Delays in project completion often lead to increased cost. Therefore, aside from increasing fleet agility, NAVFAC is interested in determining the leading factors for delays as well as predicting the projects likely to be delayed. Projects are generally split into two phases, the pre-award phase, and the post-award phase. The pre-award phase relates to the process of project submission to incorporate whether congress has provided or will provide funding as well as awarding a contract. The post-award phase details the aspects of the project once a contract is awarded. The goal of this thesis is to analyze data, using statistical analysis and machine learning algorithms to determine major factors in project delays, as well as develop predictive models for determining the likelihood of a project being delayed in the pre-award phase.

B. PURPOSE

The purpose of this research is to support NAVFAC with improving timelines for project awarding. NAVFAC has limited resources to apply to improving the current methodization for project awarding. Therefore, the goal of our research is to provide ample statistical data, and predictive models to NAVFAC for timeline improvement. We accomplish this by using statistical analysis to determine the major factors impacting project timelines. We then apply these factors to machine learning algorithms to create predictive models that will determine which projects are likely to experience delays in the awarding process.

C. RELATED WORK

In 2020, LCDR Robert Thompson conducted research into the NAVFAC project pre-award timeline, specifically the MILCON projects. His research was primarily using the random forest machine learning technique to analyze the impact of the projected initial project cost on the project awarding across multiple NAVFAC databases (Thompson, 2020). NAVFAC's annual Performance to Plan (P2P) for Military Construction outlines the previous year's analysis as well as the goal and plan moving forward for upcoming years. The FY 2021 P2P outlines the goals to develop predictive models that will flag at-risk projects for further analysis to improve and prevent delays (Komiss & Saulo, 2021).

D. ORGANIZATION

In this thesis, we focus on the collection and analysis of data available from the eProjects database. Chapter II describes the collection of data. In Chapter III, we display the information determined through statistical analysis. Chapter IV provides the methodology, model generation, and analysis of decision tree machine learning algorithms applied to the data. Chapter V provides the methodology, model generation, and analysis of the data applied to neural networks. Conclusion and recommendation for future work are provided in Chapter VI.

II. DATA INTRODUCTION

In this chapter, we cover the possible sources for data regarding NAVFAC projects. We identify which program we used for gathering data, the issues with the dataset, and the correlation with undersea warfare (USW) related projects.

A. DATASET

NAVFAC used multiple information databases to store and manage project data. Many of these databases store similar data but some are designed for tracking data in the pre-award phase, i.e., eProjects, while some are designed for the post-award phase, i.e., Electronic Construction Management System (eCMS) (Brown 2020). Below is a list of other NAVFAC databases for project management.

- Electronic Project Generator (EPG)
- eContracts
- Internet Naval Facilities Assets Data Store (iNFADS)
- Maximo

These databases tend to have overlapped data, therefore eProjects was used for the analysis of pre-awarded phase of the projects from Fiscal Year (FY) 2011–2021. Due to the redundancy and overlap of the different databases there are issues with incomplete data. Specifically, with the eProjects database there were 53240 accepted, closed-out, and completed/ready for closeout projects, but due to missing dates and cost information, only 20414 projects were able to be analyzed. As Robert Thompson discussed in his thesis in 2020, merging the data from EPG and eProjects would compile a more complete database for the pre-award phase (Thompson, 2020).

B. EPROJECTS DATABASE

eProjects is one of the major applications used by NAVFAC's Asset Management team to track projects through the pre-award phase. Our research compiled the projects from FY 2011–2021. Prior to data analysis, we removed all projects with incomplete date and cost information. While this resulted in the dataset being shrunk by approximately 60%, there was still sufficient data to analyze. Over FY 2011–2021, there were 20414 projects with complete data. For initial statistical analysis the projects were split into groups based on projected cost, shown in Figure 1. There are fewer projects with complete data for FY18–21 due to COVID-19 and delays in project completion.

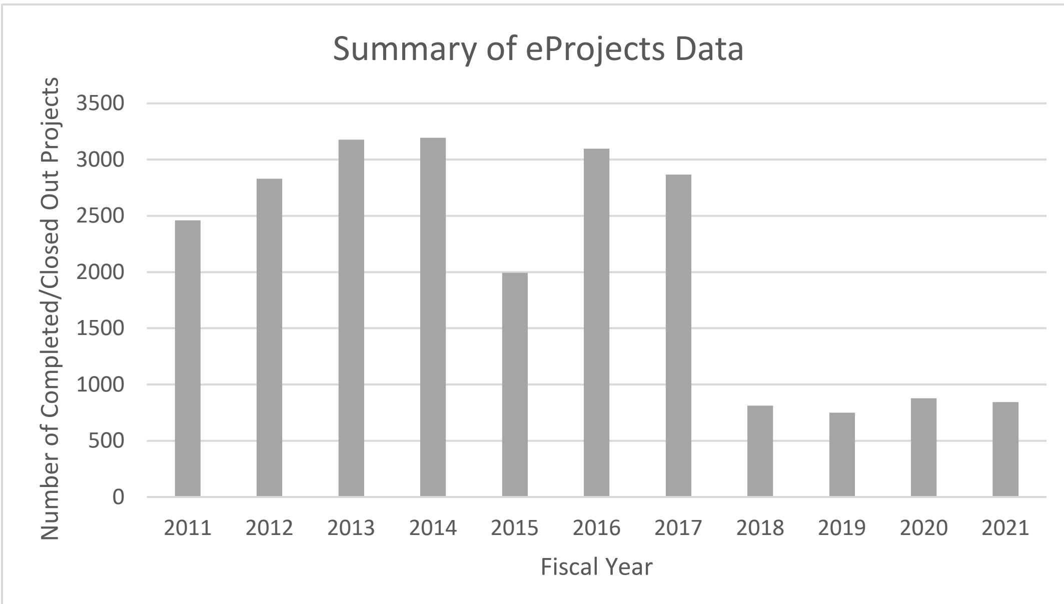

Figure 1. NAVFAC Project Distribution by FY (FY11-21)

C. INCOMPLETE DATA

As stated in section A above, 60% of the data was incomplete. Of the incomplete data most of it was missing the date of project awarding or the initial projected cost. There are two causes for the missing data. One being the responsible agencies not reporting all the information to NAVFAC, and the other being delays/failures of NAVFAC personnel to input the data into the eProjects database. Even though most of the data was incomplete, we determined that over 20000 projects were sufficient to create an accurate model for predicting whether projects would be late.

D. USW RELATED PROJECTS

Most USW projects are not processed through NAVFAC. Many projects on submarines go through Naval Support Activity Crane. Over FY11–21, the only USW related projects with complete data were through Crane. Two of the projects were for Ship Service Motor Generator work on submarines and the others were Submarine Valve Regulated Lead Acid Battery replacements. The projects submitted by Crane for submarines were awarded on-time or early.

THIS PAGE INTENTIONALLY LEFT BLANK

III. EXPLORATORY DATA ANALYSIS

In this chapter, we discuss the project parameters that we used for our statistical analysis and as our inputs for our machine learning algorithms. We show the results from the statistical analysis and the determination of parameters which have a larger impact on the project awarding process. The statistical analysis served as process for determining which parameters would serve as the fundamental inputs for the machine learning algorithms.

A. DATA ANALYSIS INTRODUCTION

Prior to applying machine learning algorithms to the dataset to develop a predictive model, we had to determine the important factors in determining the causes for a project being delayed in the pre-award phase. To accomplish this, we separated the data into multiple sections based on following parameters:

- Projected Cost
- Fiscal Year
- Month of Project Submission
- Continental United States (CONUS) verses outside the continental United States (OCONUS)
- Responsible Component
- Branch Association

Once the data was separated, we calculated the number of projects in each section, how many in each was late, and the percentage of late projects in each section.

B. PROJECTED COST

Cost plays an important role in whether a project is accepted or not due to limited budgets for new projects. During the data analysis of cost, the projects were separated into the following four categories: projected cost less than one million dollars, projected cost between one and 50 million dollars, projected cost between 50 and 100 million dollars, and projected cost greater than 100 million dollars. As shown in Figure 2, through FY11-21, 82% of projects had a projected cost less than one million dollars, and 17% of projects had a projected cost between one and 50 million dollars. For projects less than one million dollars, 36% were late; for projects between one and 50 million, 33% were late; for projects between 50 and 100 million, 28% were late; for projects over 100 million dollars, 33% were late. The small difference in late projects by initial cost indicates that the initial projected cost is not a major factor in the determination of a project being awarded late.

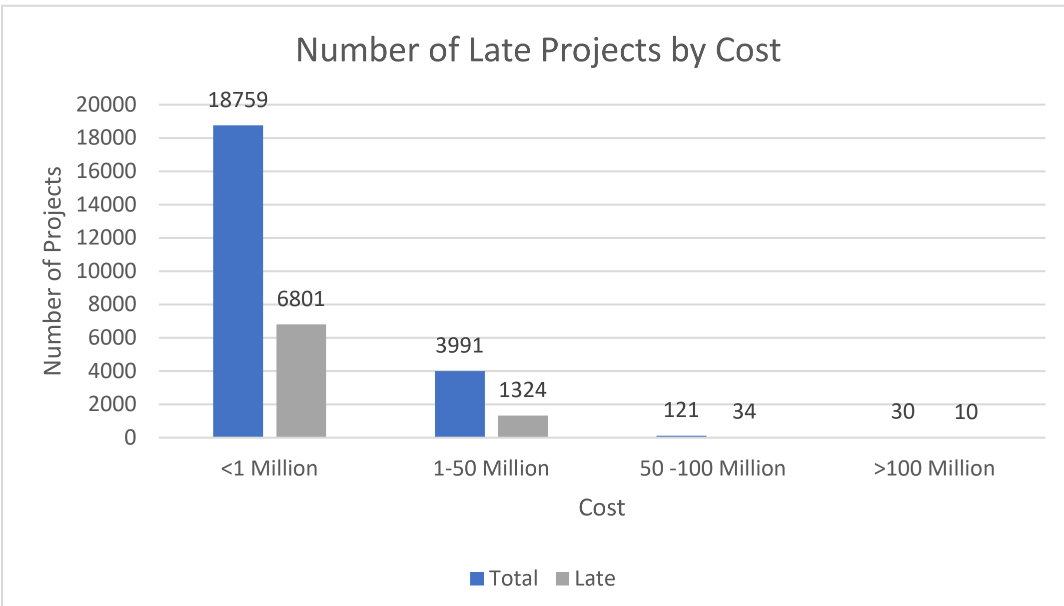

Figure 2. Number of Late Projects Separated into Projected Cost Brackets

C. FISCAL YEAR

Next, we analyzed the projects separated by FY. Through this analysis the only determination was that there were less projects in recent years, most likely due to COVID-19.

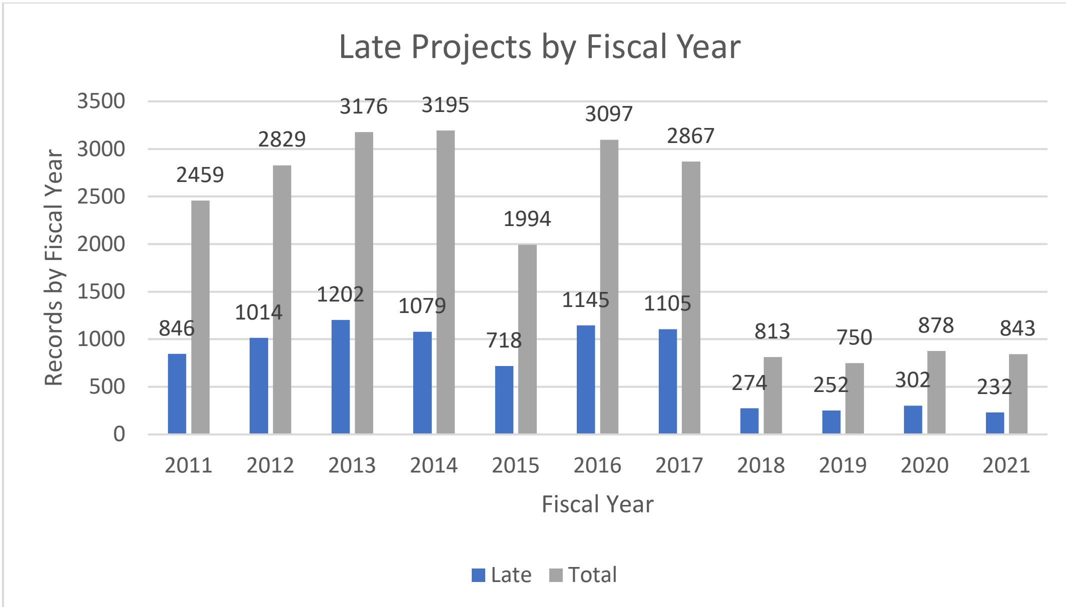

Figure 3. Number of Projects and Late Projects Separated by FY

Figures 3 and 4 show that the large disparity between the number of projects in the earlier FYs and the more recent years still did not have an impact on the percentage of projects that were late in the awarding process. Therefore, with a negligible difference in the percentage of late projects across FYs we decided not to include FY in the machine learning algorithms for predictive model generation.

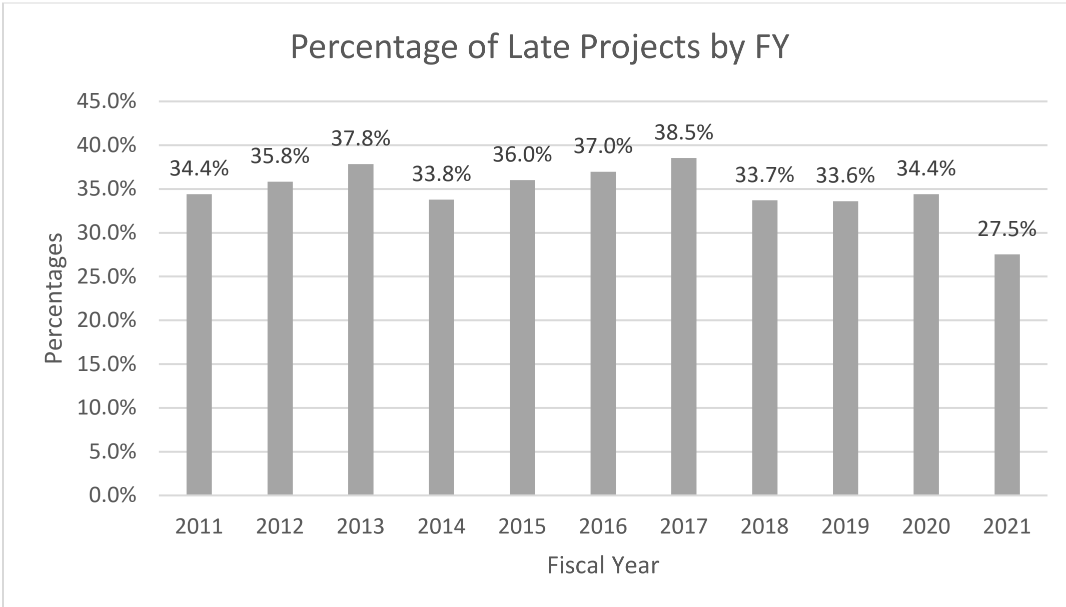

Figure 4. Percentage of Projects Late for Award by Separated by FY

D. MONTH OF PROJECT SUBMISSION

We analyzed when the projects were submitted during the FY due to the additional time constraints for project approval. Aside from the projected cost of the project, this was determined to be one of the leading causes for delays in project approval. Through the analysis, shown in Figure 5, we determined that projects submitted closer to the end of the FY had a higher chance of being approved on time or early.

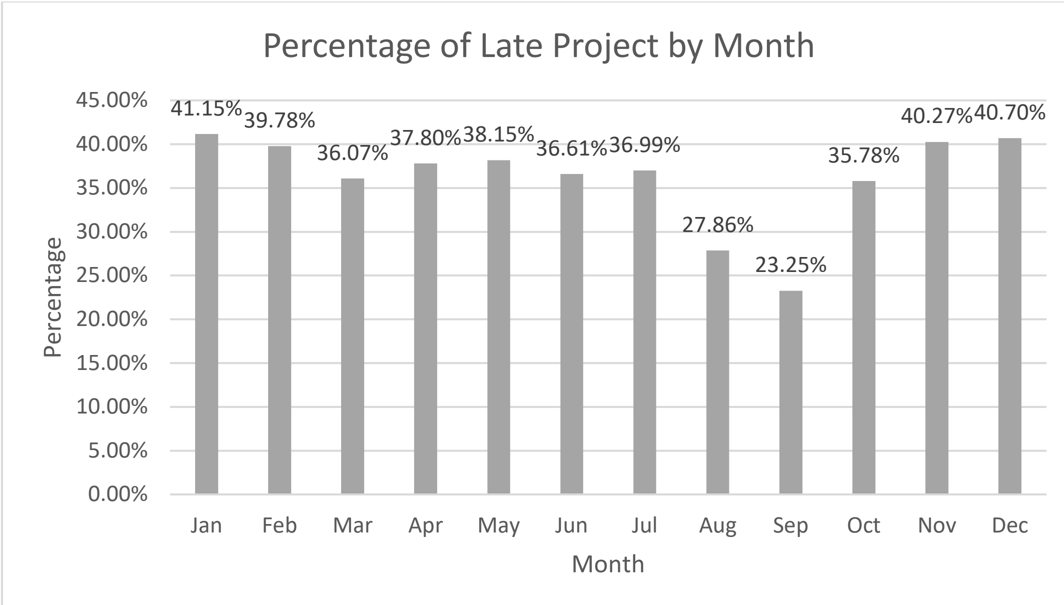

Figure 5. Percentage of Late Approvals Separated by Month of Project Submission

After noting that projects submitted in August and September had a much lower percentage of late awards, we questioned whether there were less projects being submitted, leading to a better approval rate. Based on the number of projects submitted each month, that could not be the case. August had the second most projects submitted and the second lowest percentage of late awards, while September had fewer projects submitted than most months, it also had the lowest percentage of late projects (see Figure 6).

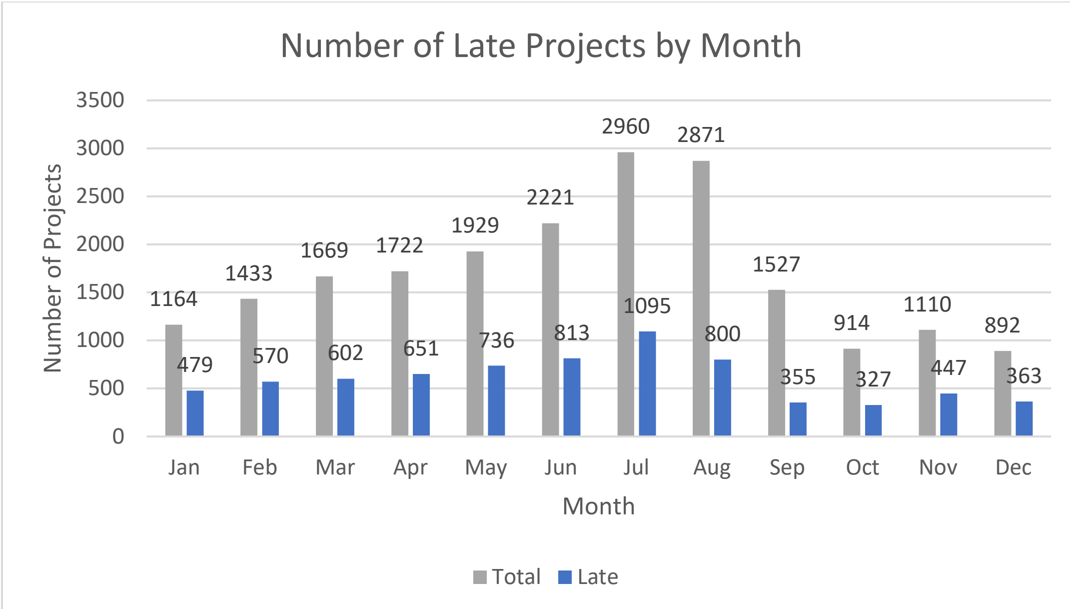

Figure 6. Number of Projects and Late Awards Separated by Month of Project Submission

E. CONUS/OCONUS

Next, we thought that whether the project was submitted for CONUS would have a higher on time approval rate than projects submitted for overseas. Figure 7 displays that 60% of projects are CONUS while the other 40% are OCONUS.

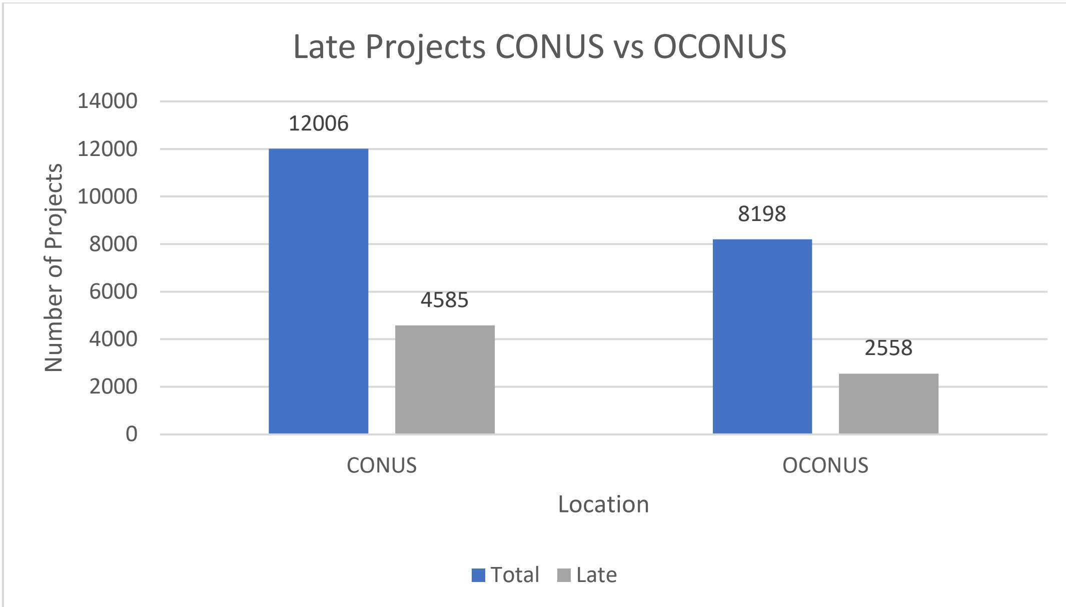

Figure 7. Total Number of Projects and Late Projects Separated by Location

Through calculations we determined that OCONUS projects had a better percentage of approval time (see Figure 8). The difference between the percentage of late projects whether they are CONUS or OCONUS made it one of the factors entered into the machine learning algorithms.

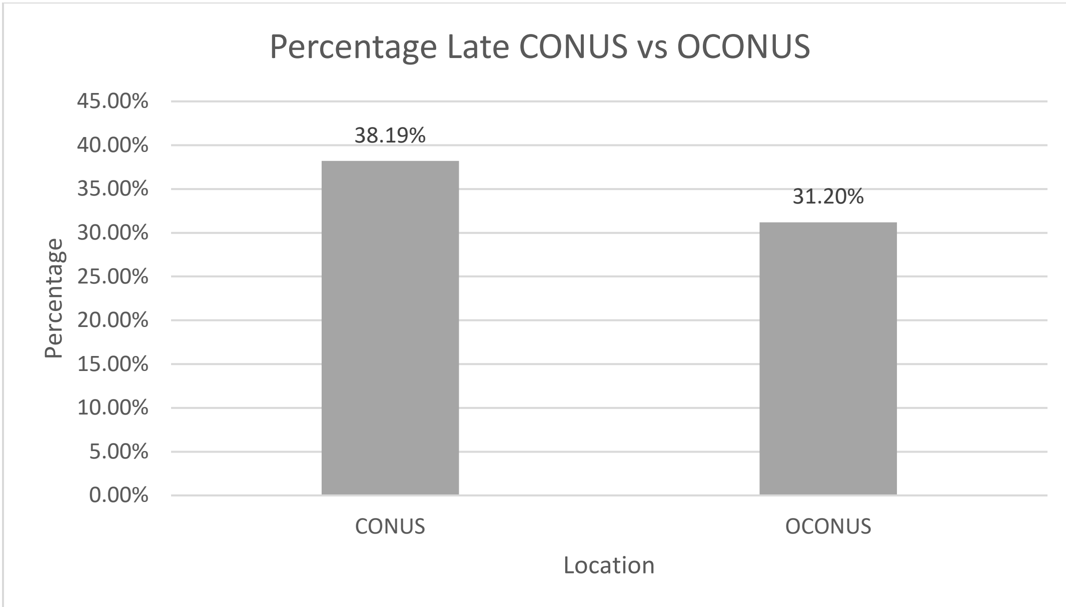

Figure 8. Percentage of Late Projects Separated by Location

F. RESPONSIBLE COMPONENT

One of the other data points in the eProjects database is the responsible component. The responsible component is the agency that is responsible for the submission of the project. The agency that submits the project is responsible for determining the projected cost, when the project is submitted, and the wording/reason for the project. Therefore, the agency responsible for the project plays an important part in the timeliness of project award. If the reason for the project is well-defined, the project has a higher chance of being awarded on time due to the overall necessity of the project. Figure 9 shows that there is a large disparity between the number of projects submitted by different agencies.

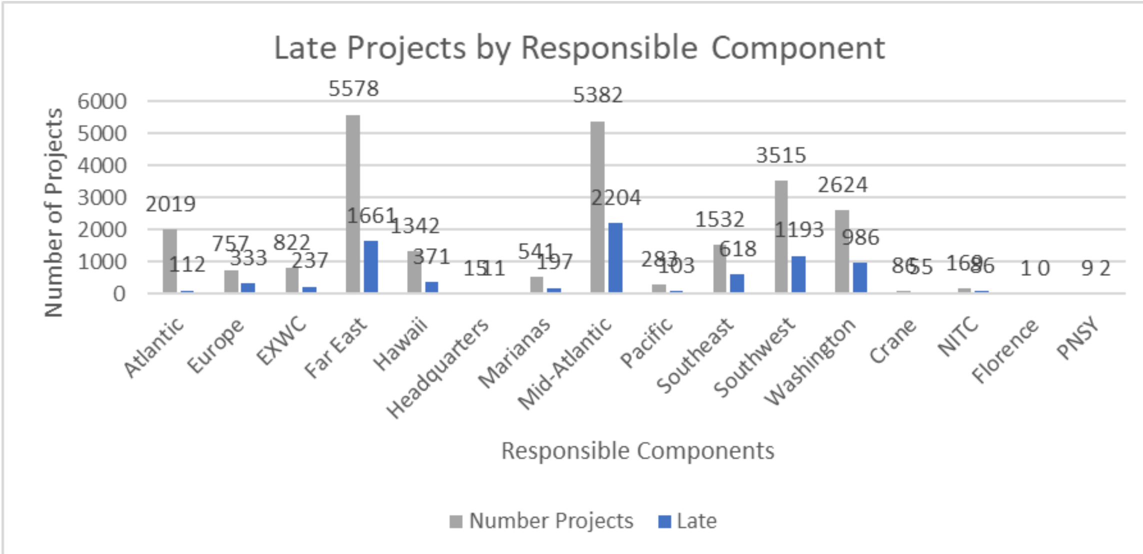

Figure 9. Number of Projects Submitted Separated by Responsible Component

Most of the responsible components shown above are different branches of NAVFAC. Even though the projects are submitted through NAVFAC agencies, in some cases, they have vastly different rates at which projects are awarded on time. Shown in Figure 10, the NAVFAC Headquarters, while only submitted 15 projects over the past decade, had over 70% late in the awarding process. This is similar for Crane (designs the batteries and Ship Service Motor Generators for the submarine force), who submitted 86 projects in the last decade with 64% of them being late for awarding. The large disparity for percentage of late projects between the different responsible components, implies that the quality of the project submittal report could play a factor in the whether the project is awarded on time.

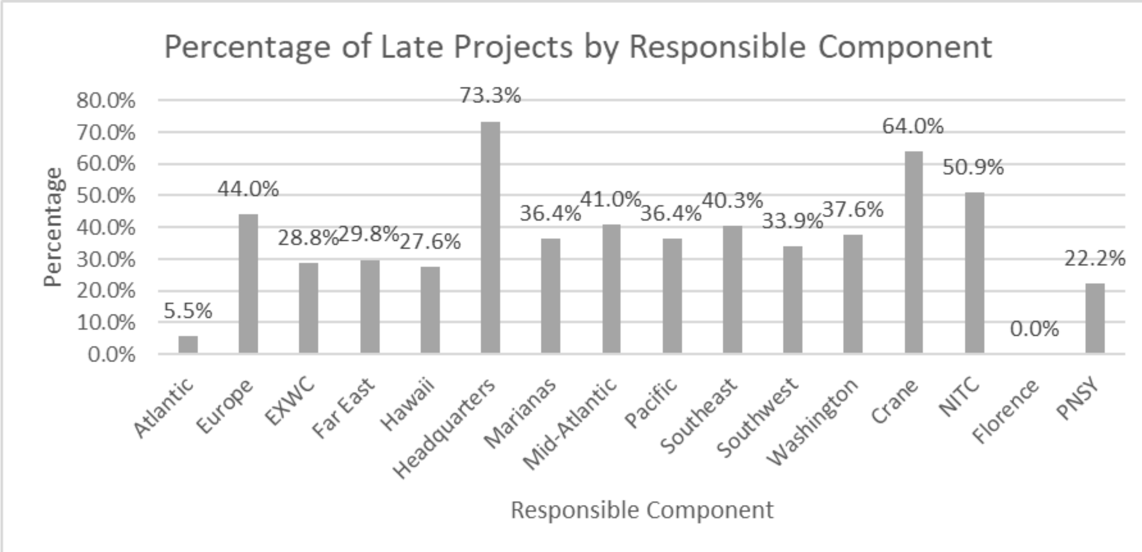

Figure 10. Percentage of Late Projects Separated by Responsible Component

G. BRANCH ASSOCIATION

The last data point we analyzed in Figure 11 was the branches associated with the project. Through analysis we split the data into five components for branch association. There are projects that are associated with the Army, Air Force, Marine Corps, more than two branches, and just the Navy. As expected, most projects have no associated branch other than the Navy. The Marine Corps has many projects in conjunction with the Navy as well, but there are very few projects associated with solely the Army and Air Force.

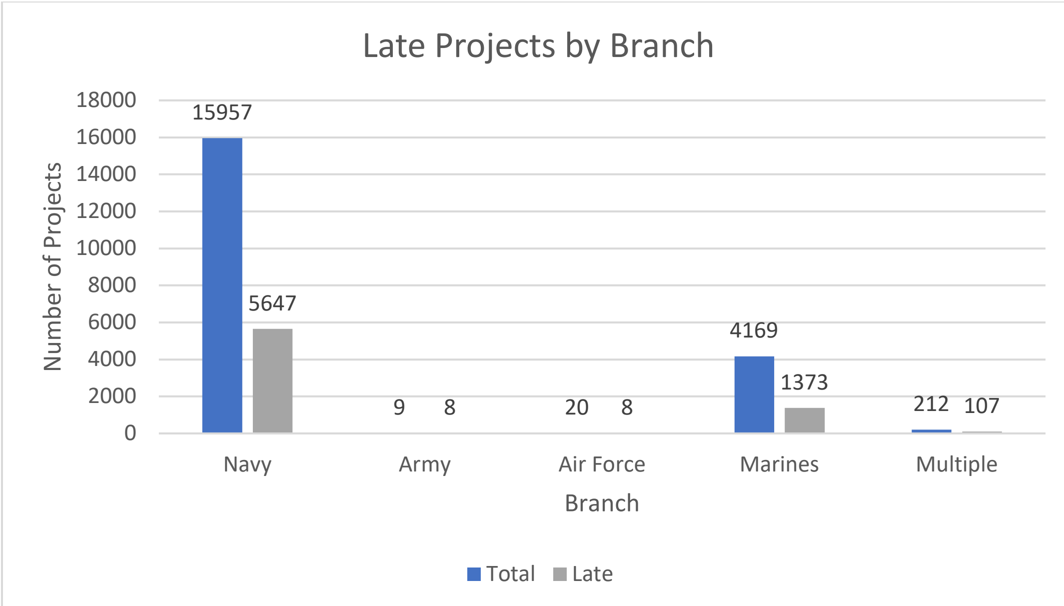

Figure 11. Total Projects Separated by Branch Association

Even though there are few projects associated with the Army, Air Force, and multiple other branches, they all have a higher percentage of late projects than those solely submitted by the Navy, and those associated with the Marine Corps (Figure 12). Due to the low number of projects associated with branches other than the Marine Corps, we decided there was not sufficient data for machine learning algorithms to determine the impact of branch associated accurately.

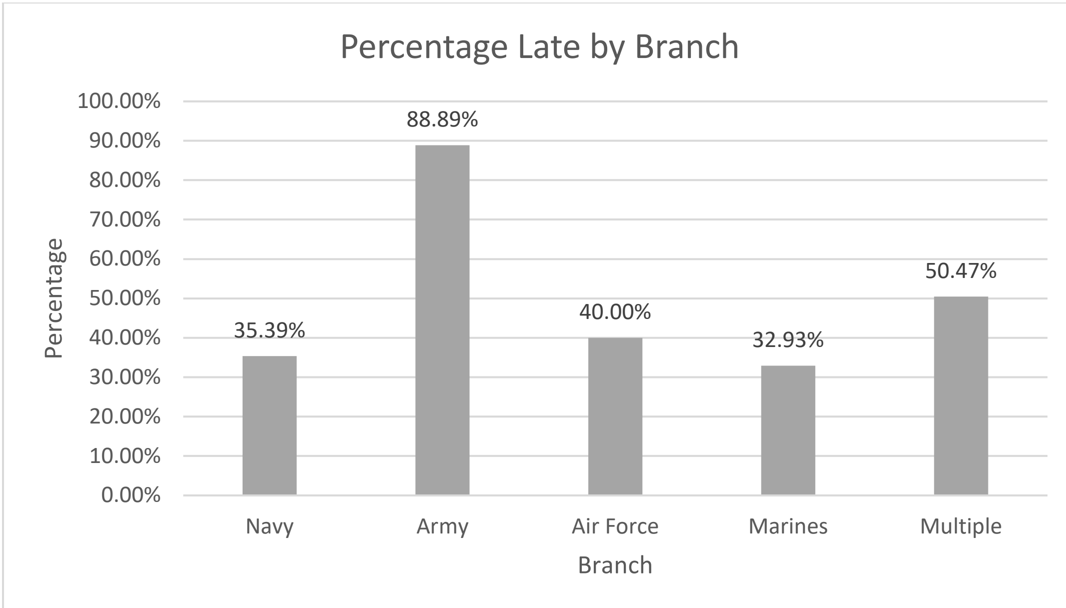

Figure 12. Percentage of Late Projects Separated by Branch Association

Due to the low number of projects associated with the Army and the Air Force, we analyzed the data by moving the Army and Air Force associated projects to the Multiple. The low number of projects would possibly negatively impact the generated models. The results of the data adjustment are shown in Figures 13 and 14 and indicate that if a project is associated with any branch other than the Marine Corps and Navy it has a greater than 50% chance of being awarded late.

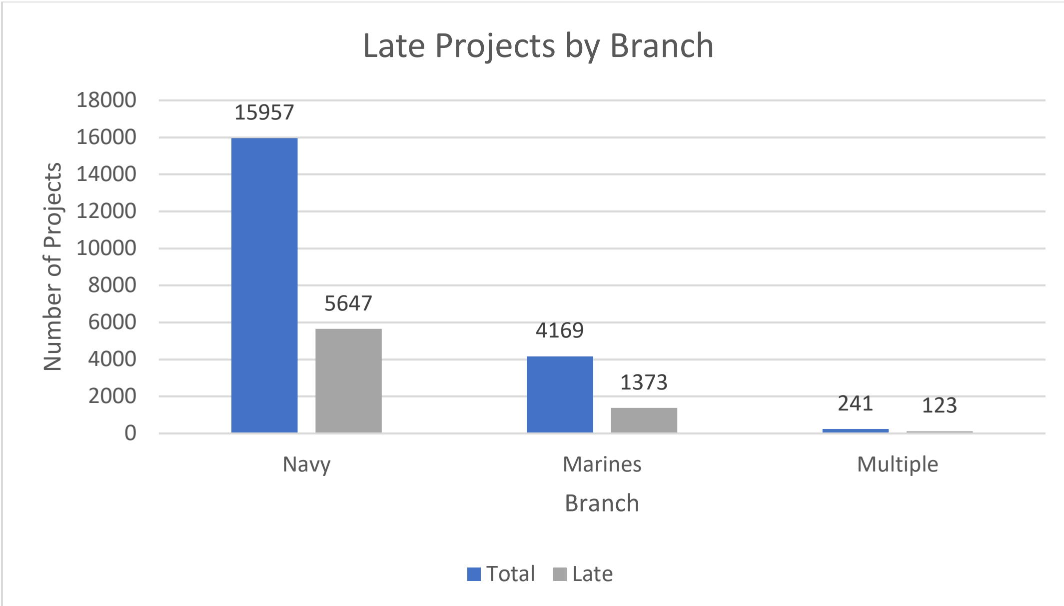

Figure 13. Number of Projects and Late Projects by Branch Association

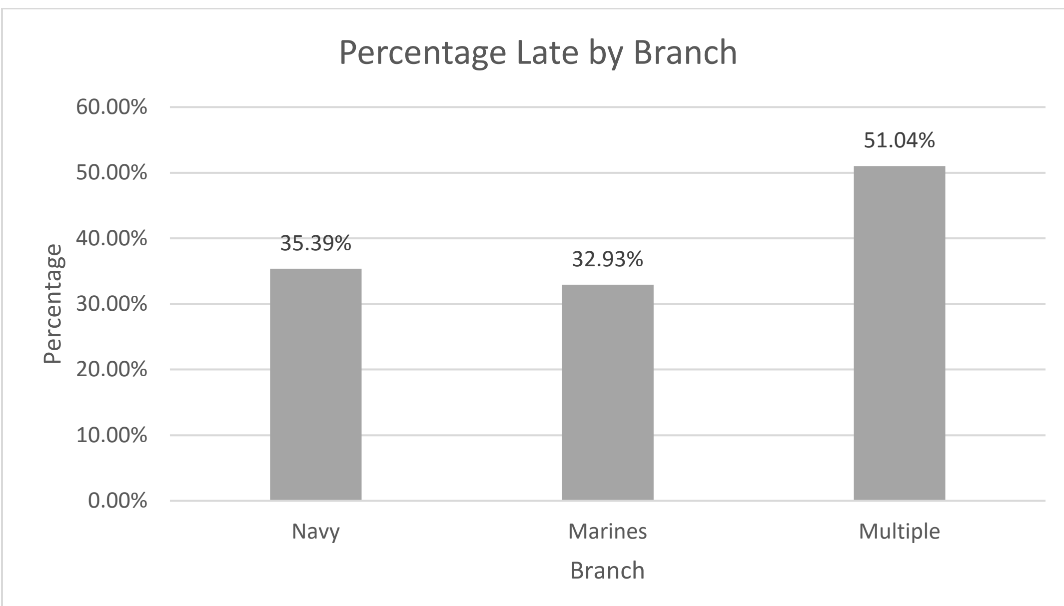

Figure 14. Percentage of Late Projects by Branch Association

THIS PAGE INTENTIONALLY LEFT BLANK

IV. MACHINE LEARNING—DECISION TREES

In this chapter, we cover the basics of decision trees, what decision trees we used, the trained models, and the results of the models. The chapter concludes with the accuracy of the trained decision tree models.

A. DECISION TREES

Decision Trees are constructed with directed graphs (Kamiński et al., 2017). Most of these graphs are weakly connected as they only allow traversal in one direction. The trees are created by denoting probabilities and payoffs. An example of a decision tree for our analysis of NAVFAC project data is shown in Figure 15.

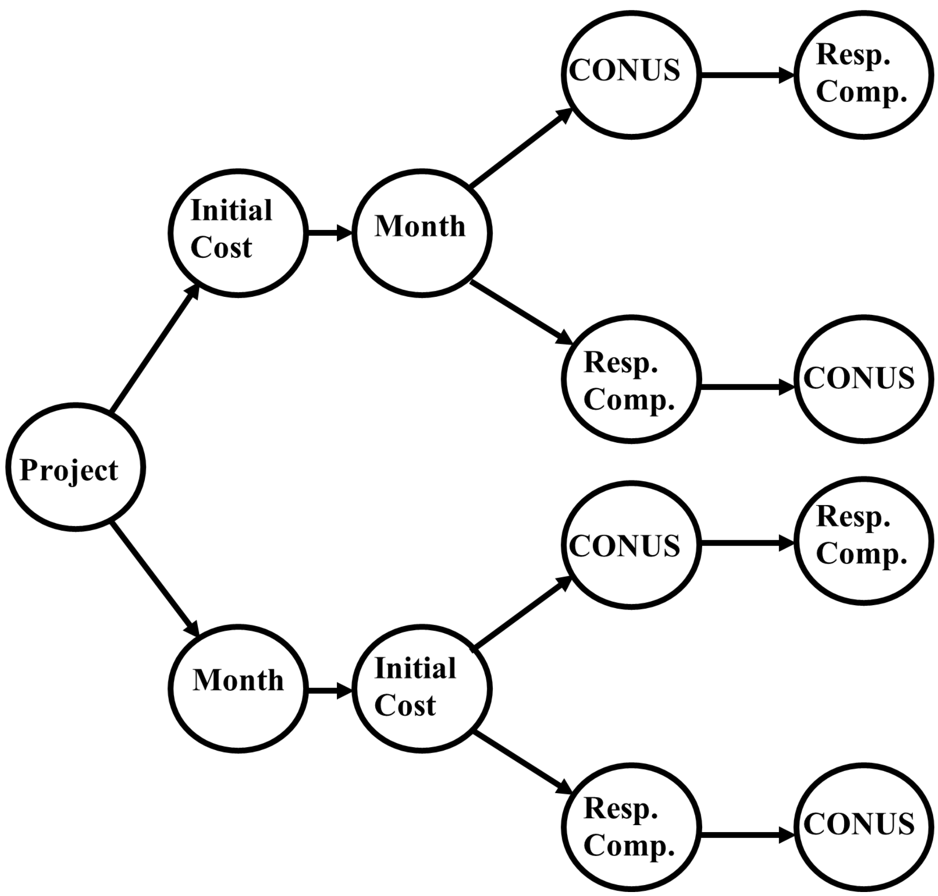

Figure 15. A Sample Decision Tree

We utilized MATLAB's Machine Learning toolbox to train and analyze the data. We trained on the data from spanning FY11–20, saving the FY21 to test the accuracy of the generated models. MATLAB's machine learning toolbox includes five different decision tree models: Fine Tree, Coarse Tree, Medium Tree, Bagged Tree, and Boosted Tree. To create the most accurate model, we trained the data to each of the decision trees as well as using different combinations of data inputs.

In 2020, Robert Thompson used the random forest machine learning technique to analyze the project cost on whether projects were late (Thompson, 2020). We used the machine learning application in MATLAB. For continuity between theses, we applied the bagged tree machine learning algorithm as it uses the random forest algorithm.

B. MODEL TRAINING

Figures 16, 17, and 18 show the trained predictive models for Fine Tree, Medium Tree, and Coarse Tree. The difference among these three models is the Coarse Tree model has a maximum of four splits, the Medium Tree has a maximum of 20 splits, and the Fine Tree has a maximum of 100 splits (Mathworks, Decision Trees, 2022). Each Tree was trained using all the following parameters from the data set: Initial cost, month of application, responsible component, CONUS/OCONUS, and associated branch.

Figure 16. Graphical Representation of Fine Tree Model (40 Splits)

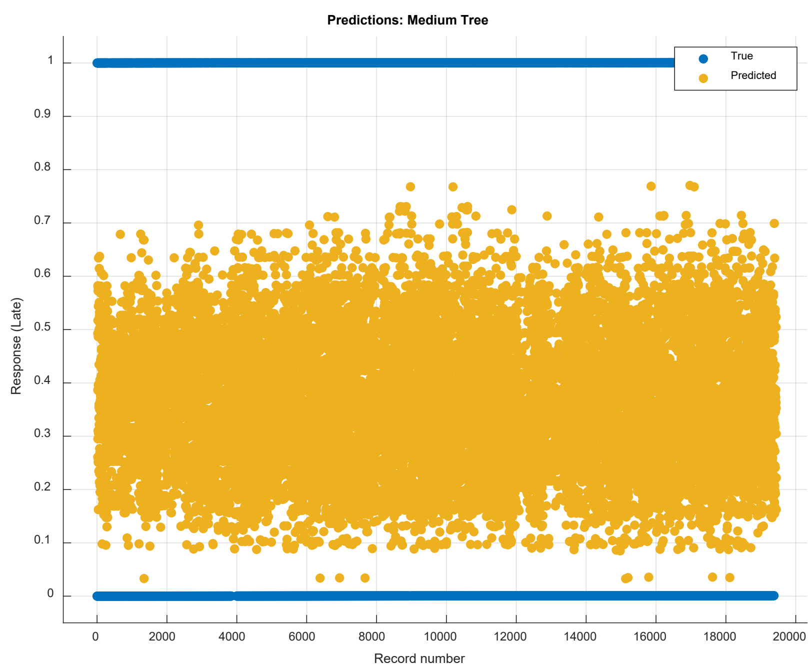

Figure 17. Graphical Representation of Medium Tree Model (20 Splits)

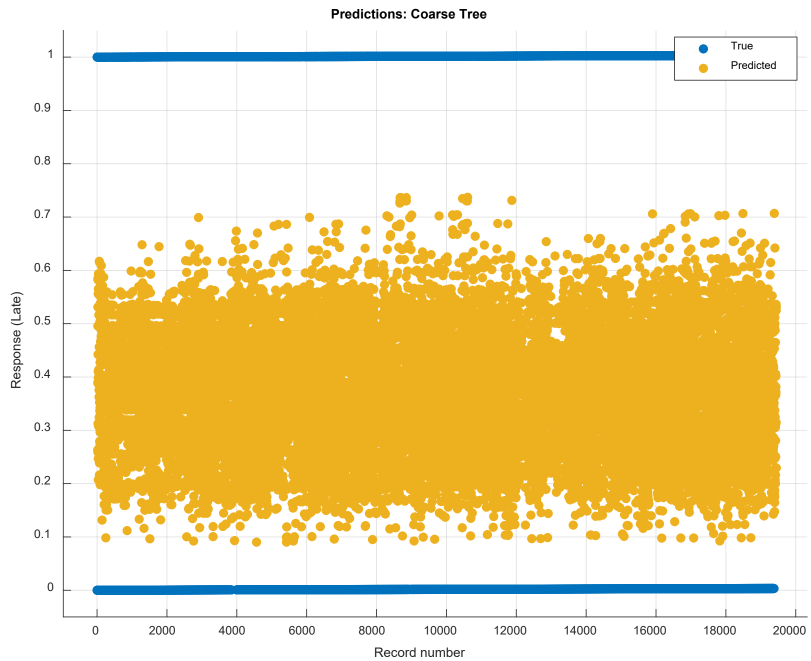

Figure 18. Graphical Representation of Coarse Tree Model (4 Splits)

Figures 19 and 20, show that with fewer splits in the decision trees the data is more centralized around the statistical analysis results between 30–50%, whereas the Fine Tree has results spread across between 0–100%. The Bagged Trees model uses a random forest machine learning method with a high number of learners (Mathworks Inc., Decision Trees, 2022). Figure 20 shows the data for the Bagged Tree model being more spread across the spectrum of 0–100 as well most likely due to the maximum number of learners being set to 100 to match the Fine Tree Model. Boosted Trees use the standard decision tree model with AdaBoost included. AdaBoost is an iterative process in which the machine learning algorithm turns weak classifiers into strong classifiers by learning from the mistakes (Kurama, 2021). The Boosted Tree model, Figure 21, show results centered around 30% chance of the projects being late for award. For the Boosted Tree Model, we set the maximum learners to 100 to match the capabilities of the Fine Tree model due to the high performance of the Fine Tree Model. We set the minimum leaf size to eight as changing the minimum leaf size did not have an appreciable impact on the performance of the model.

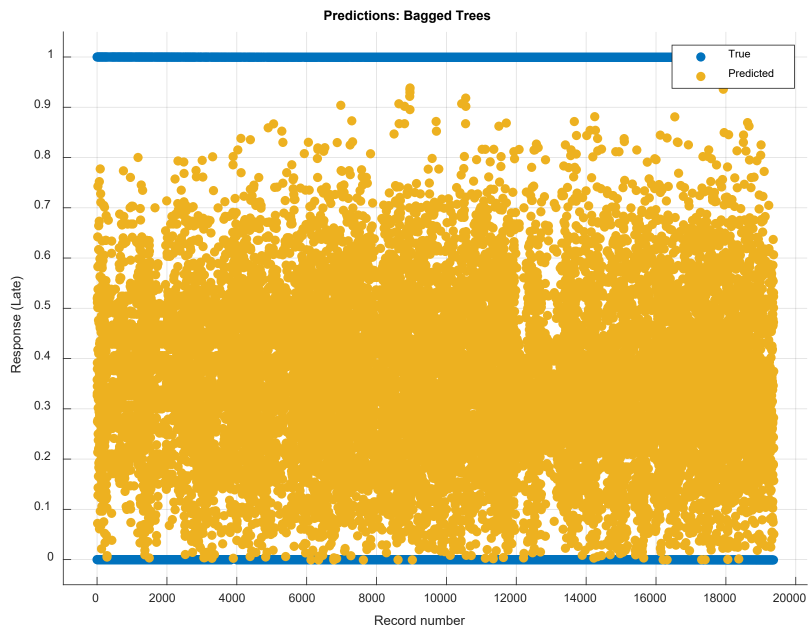

Figure 19. Graphical Representation of Bagged Tree Model

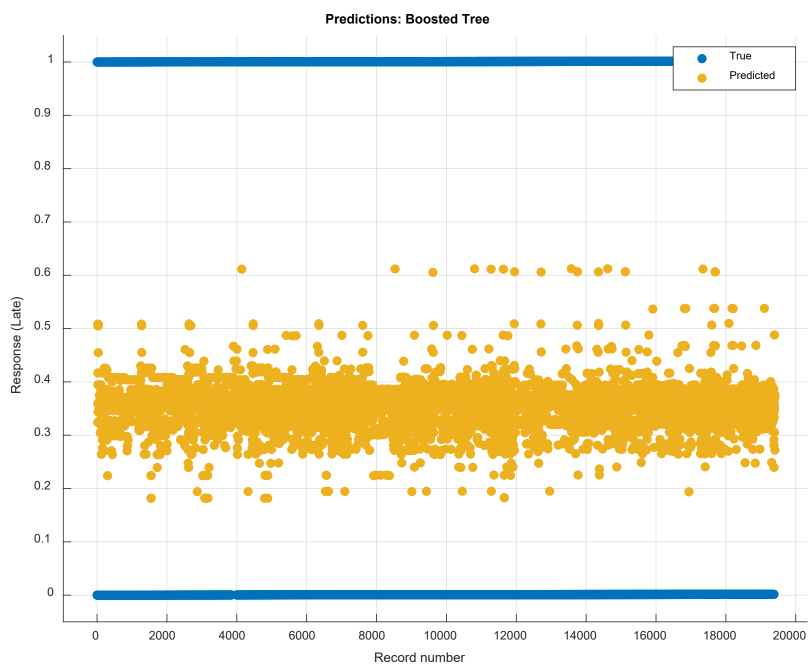

Figure 20. Graphical Representation of Boosted Tree Model

While the above models were trained using five parameters, we also trained the models three other times using different numbers of parameters. We trained the model without including the branch, without including the responsible component, and without the branch and responsible component. All models were similar to the Figures 17–21 with slight variances.

C. MODEL TESTING

As discussed in the sections above, we applied the trained models to FY21 to test the accuracy. The main difference between the trained figures and the testing figures is that the testing was done on 842 projects while we trained to 20000 projects. Figures 21–25 show graphical representation of the test results for the models with all five parameters used for training. Each of the test models show similar representation with the test results as they did with the trained models.

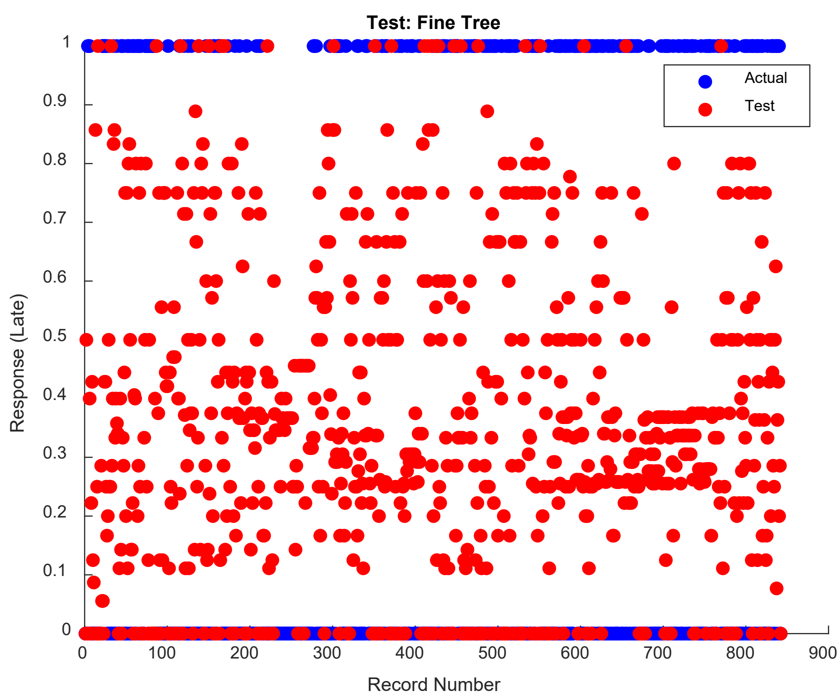

Figure 21. Test Results for Fine Tree Model (40 Splits)

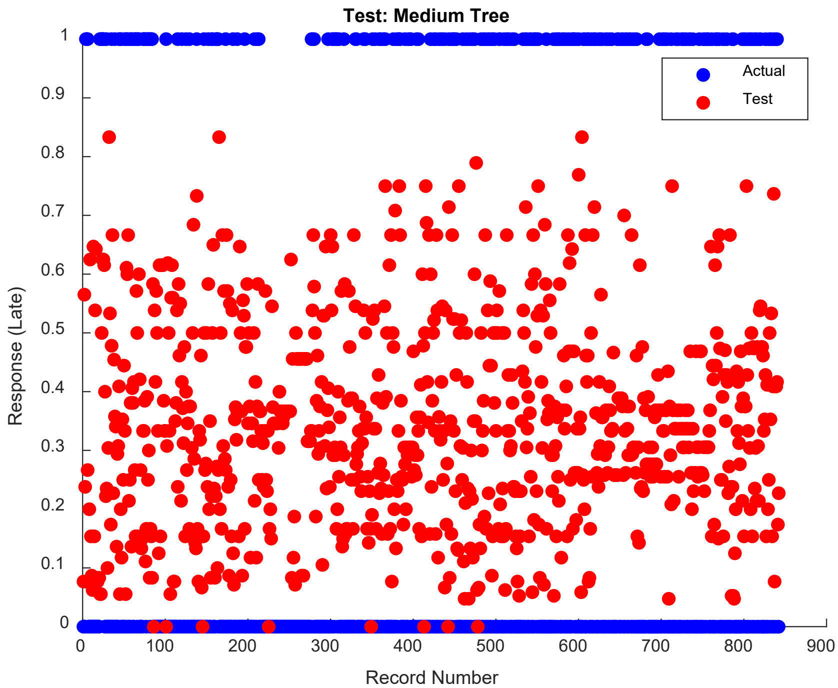

Figure 22. Test Results for Medium Tree Model (20 Splits)

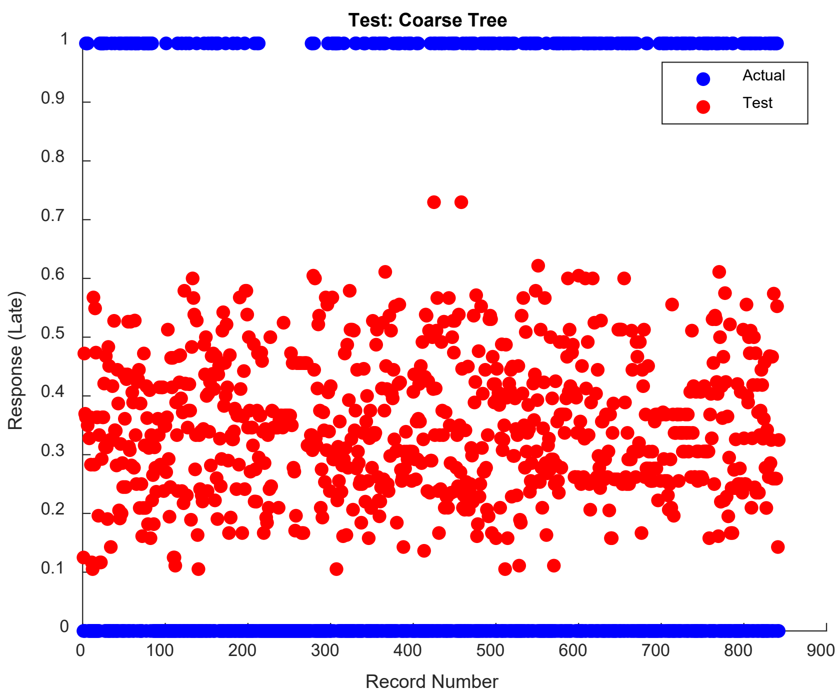

Figure 23. Test Results for Coarse Tree Model (4 Splits)

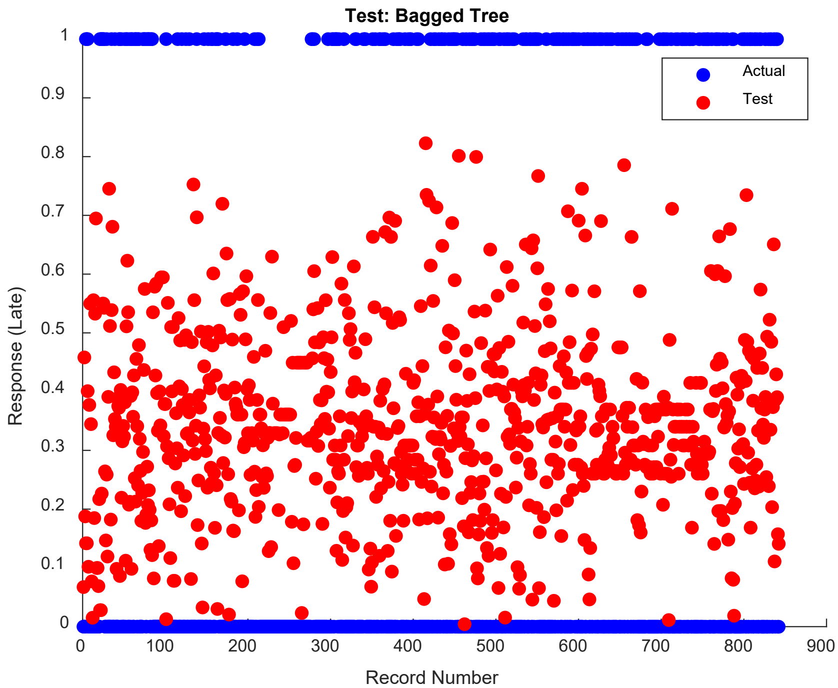

Figure 24. Test Results for Bagged Tree Model

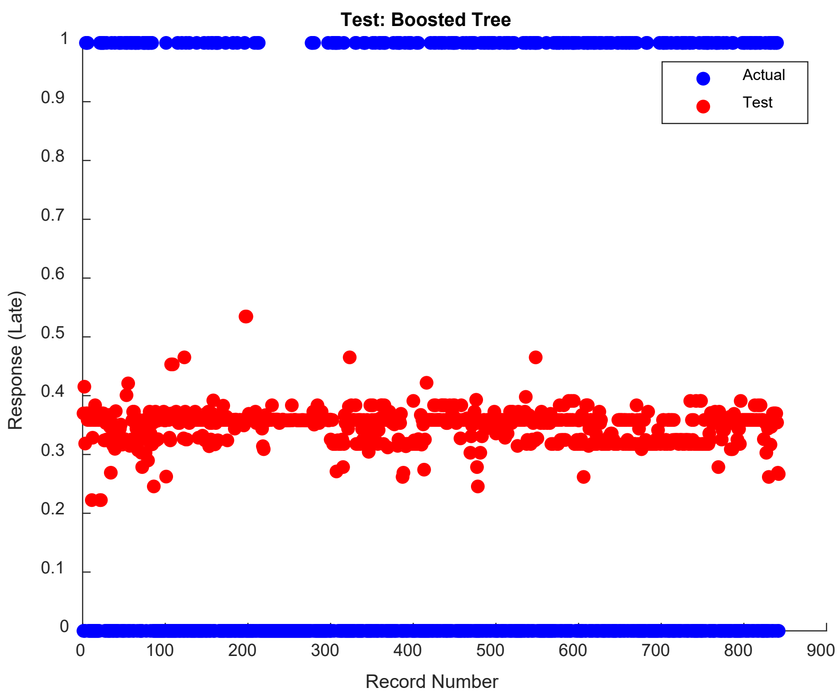

Figure 25. Test Results for Boosted Tree Model

Through model training and testing the Fine Tree model was the only model that showed any projects having a 0% and a 100% chance of being award late. The Boosted Tree model seems the least effective with almost all projects having around a 35% chance of being awarded late.

D. MODEL ACCURACY

To further test the accuracy of the model, we adjust the test results to have a resulting value greater than or equal 0.5 to be set to one, and any value less than 0.5 to be set to zero. This adjusted the result to only have outputs of zero and one, and therefore the results could be directly compared to the actual results from FY21. The results are shown in Table 1. As predicted by the trained models, the Fine Tree model was the most accurate model ranging between 92–96% accuracy. The Boosted Tree model was the least accurate with an accuracy of 72%. The inaccuracy of the boosted tree model is most likely caused by the hyperparameters used for boosting. With AdaBoost the Boosted Tree model acts similar to a neural network and requires a larger dataset to create an accurate predictive model. The boosted tree model does not provide any useful data as it only predicted a few projects as being late resulting in the 72% accuracy (approximately 27% of projects were late in FY21).

Table 1. Decision Tree Test Accuracy (FY21)

ALL PARAMATERS	ACCURACY	NO BRANCH	ACCURACY	NO RESPONSIBLE COMPONENT	ACCURACY	NO BRANCH OR RESPONSIBLE COMPONENT	ACCURACY
Fine Tree	94.54%	Fine Tree	95.37%	Fine Tree	95.49%	Fine Tree	92.76%
Medium Tree	88.24%	Medium Tree	88.48%	Medium Tree	90.02%	Medium Tree	87.89%
Coarse Tree	84.09%	Coarse Tree	83.37%	Coarse Tree	83.13%	Coarse Tree	80.76%
Bagged Trees	86.10%	Bagged Trees	85.75%	Bagged Trees	89.19%	Bagged Trees	86.94%
Boosted Tree	72.68%	Boosted Tree	72.92%	Boosted Tree	72.68%	Boosted Tree	72.57%

The confusion matrix, Figure 26, shows that the Fine Tree model is predominantly true positives and true negatives, confirming the accuracy of the model. There are still a significant number of false positives and negatives showing that there are improvements that can be made to the mode.

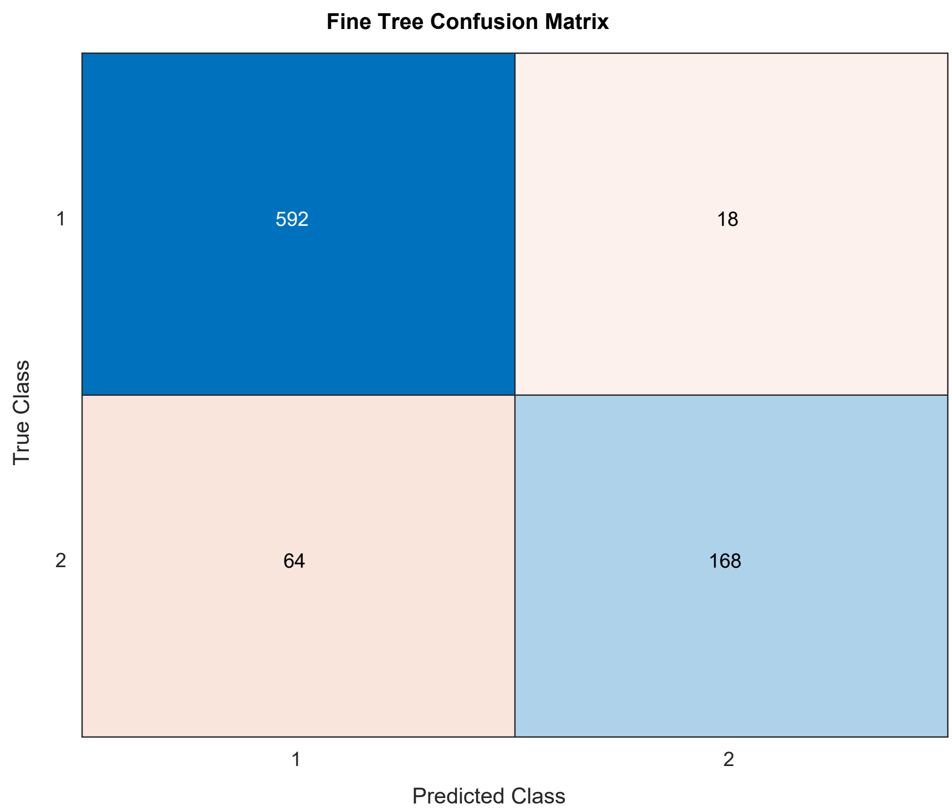

Figure 26. Confusion Matrix for Fine Tree Model Results

V. MACHINE LEARNING—NEURAL NETWORKS

In this chapter we cover a brief overview of neural networks, followed by options for neural networks in MATLAB's regression learner application. We then discuss the methods used for training models along with the models trained by each neural network available. This chapter concludes with our testing of the models and the results of the tests.

A. NEURAL NETWORKS

Neural networks are adaptive systems made up of interconnected nodes and layers. Neural networks can learn from inputted data. To train the gathered data, we used the same process as with the decision trees, using FY11–20 to train the data, and then used FY21 to test the trained data.

Figure 27 displays an example neural network with three total layers. The initial layer in a neural network is the input layer. The inputted data is then weighted and applied to the hidden layer. This process is continued for each of the hidden layers until the output layered is reached. Figure 26 shows the calculation process that occurs in each of the hidden layers and the output layer for any neural network. These calculations determine each of the individual nodes of each layer to determine the overall output. Equations (1), (2), and (3) show the different activation functions available in MATLAB's machine learning toolbox. The use of these activation functions can be seen in calculation process shown in Figure 28 (Goodfellow, 2017).

$$Rectified\ Linear{:}\quad f(z) = \begin{cases} 1\ if\ z > 0 \\ 0\ if\ z \leq 0 \end{cases} \tag{1}$$

$$Hyperbolic\ Tangent{:}\quad f(z) = \frac{e^{z} - e^{-z}}{e^{z} + e^{-z}} \tag{2}$$

$$Sigmoid{:}\quad f(z) = \frac{e^{z}}{1 + e^{z}} = \frac{1}{e^{-z} + 1} \tag{3}$$

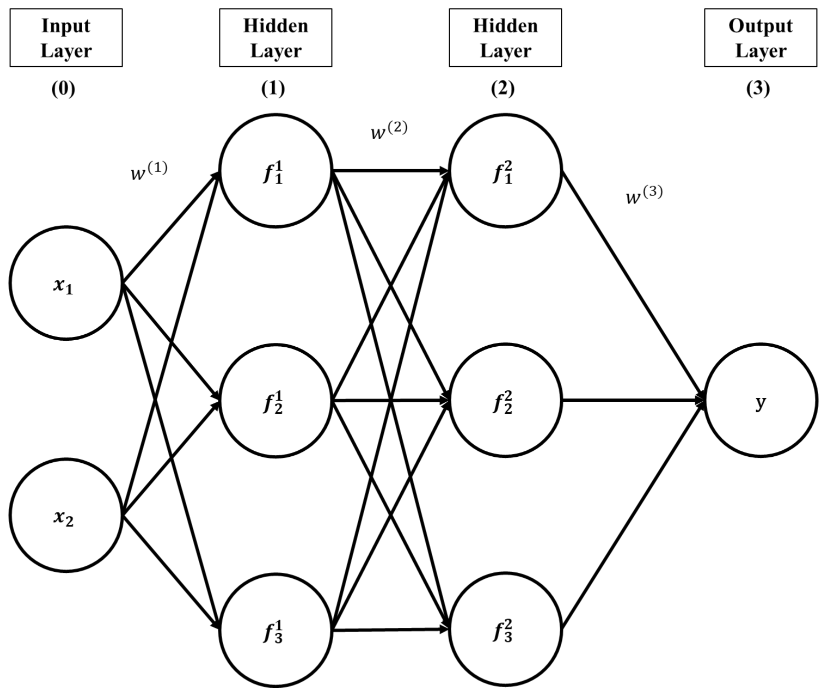

Figure 27. A Sample Neural Network

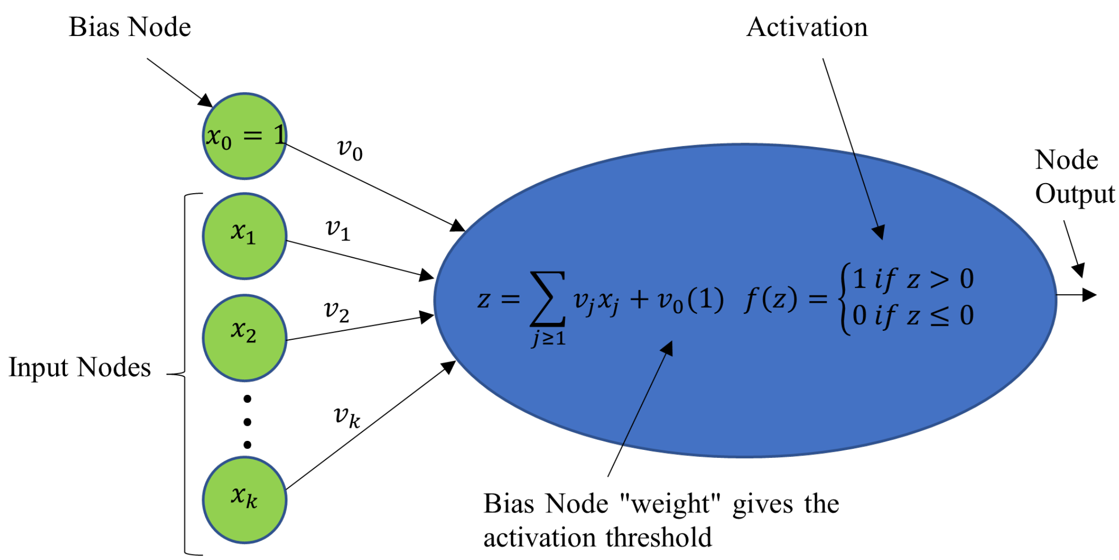

Figure 28. Calculation Process for Neural Networks

MATLAB's machine learning toolbox has five built-in neural networks that we used for modeling: narrow, medium, wide, bilayered, and trilayered. The differences between the narrow, medium, and wide neural networks are the learning flexibility based on the size of the first layer, while the bilayered neural network flexibility is based on the first and second layers, and the trilayered neural networks is based on the first three layers (Mathworks Inc., *Choose Regression Model Options* 2022). MATLAB's machine learning toolbox also has limitations. The regression learner application we used for data analysis and model generation limits the neural networks to three layers maximum and does not allow the user to adjust of weights between layers. The regression learner also allows for choosing of the activation function. The options for activation function are, Rectified Linear Unit (ReLU), sigmoid, and hyperbolic tangent (tanh) (Mathworks Inc., *Regression Learner Application* 2022). We tested models using each of the three activation functions.

B. MODEL TRAINING

Figures 29, 30, and 31 are the trained models for narrow, medium, and wide neural network models representing the FY11–FY20. These models are three layered neural networks with 50 nodes in the first layer, 25 nodes in the second layer, and 10 nodes in the third layer. We varied the number of nodes in first layer starting with 10 nodes and increasing by 10 up to a final value of 100. The number of nodes in the second and third layers were increased starting by 5 up to 50. Changes in the size of the nodes did not change the overall accuracy of the model in an appreciable manner. We also varied the number of layers from one layer to three layers with similar negligible changes in the model accuracy. The activation function used for the neural networks did have a large effect on the resulting models. Testing each of the activation functions resulted in tanh providing the most accurate models. Therefore, for the final model training we determined that using the three layers discussed above with tanh activation function was optimal for maintaining similarities for comparing the different neural networks while meeting the minimum requirements for each of the neural network modeling requirements.

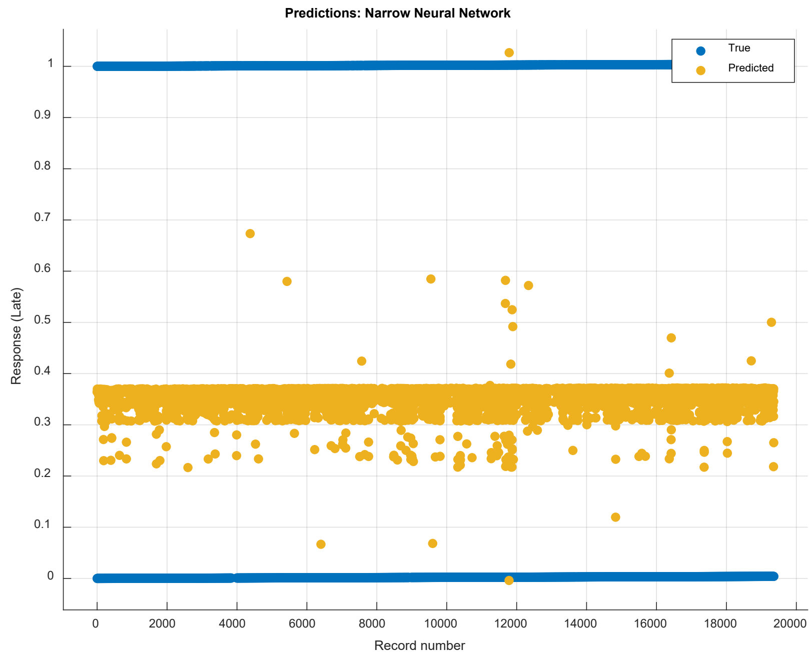

Figure 29. Graphical Representation of Narrow Neural Network Model

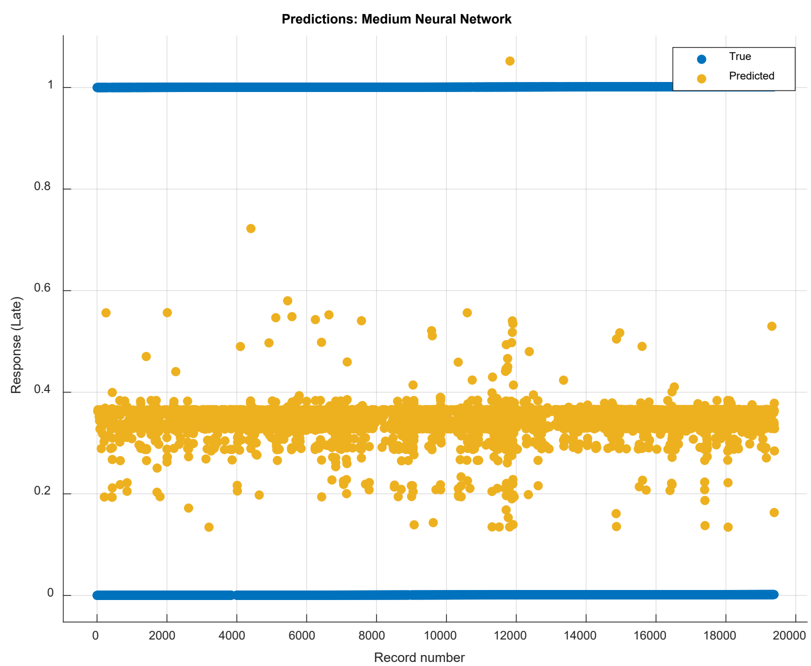

Figure 30. Graphical Representation of Medium Neural Network Model

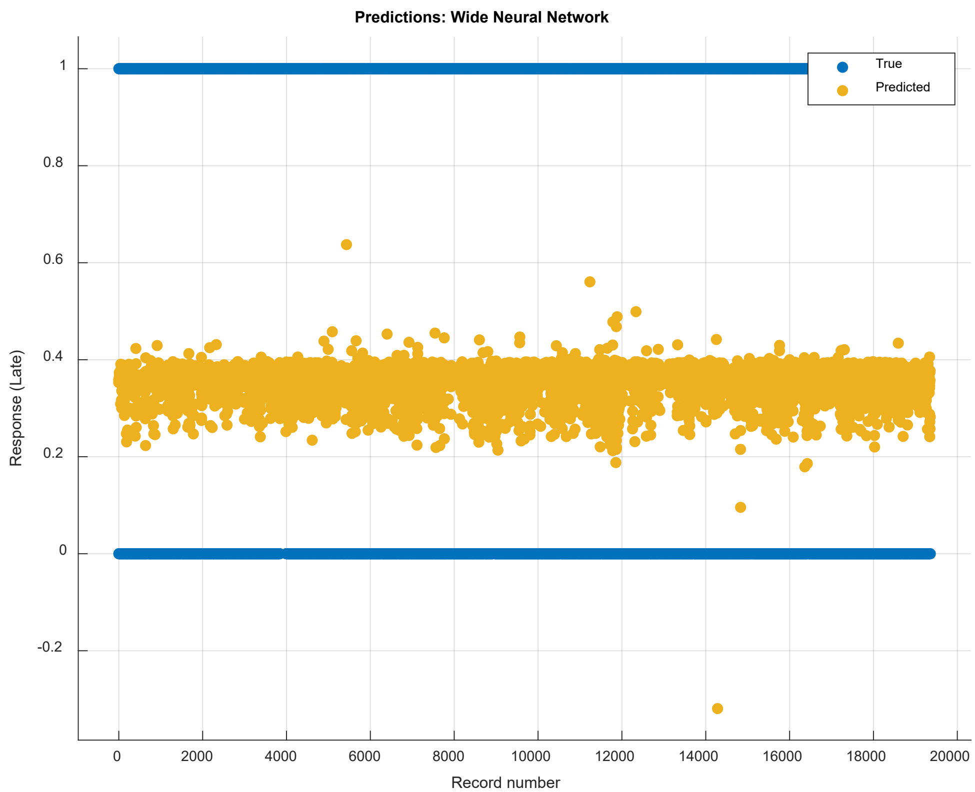

Figure 31. Graphical Representation of Wide Neural Network Model

Each of the neural network models shown above are centered between 30–40% with the narrow neural network having the tightest grouping around 35%. For model testing purposes, this shows that the models will not be as accurate as the decision trees due to using 50% as the deciding factor for whether a project will be late. The bilayered and trilayered neural network models, shown in Figures 32 and 33, have similar results with model predictions centered between 30–40% with few projects having greater than a 50% of being awarded late.

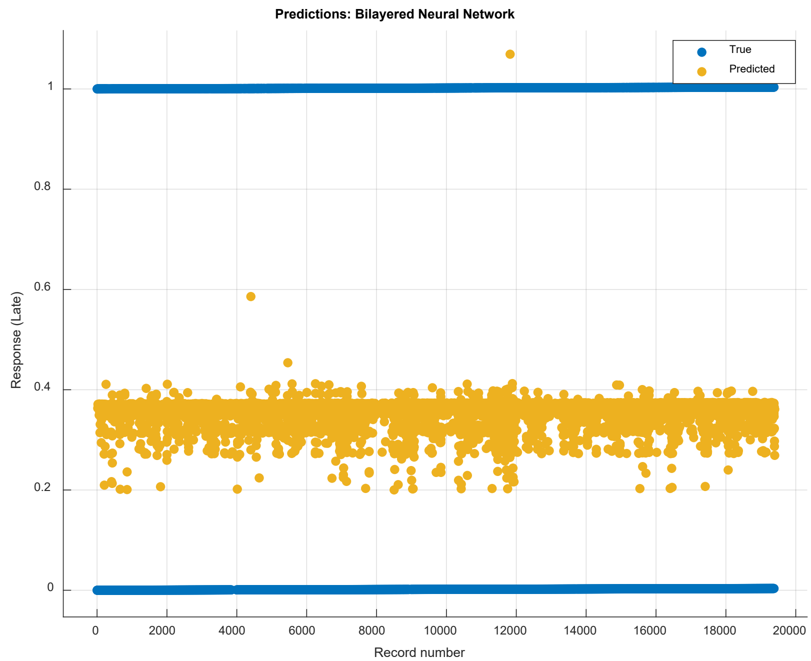

Figure 32. Graphical Representation of Bilayered Neural Network Model

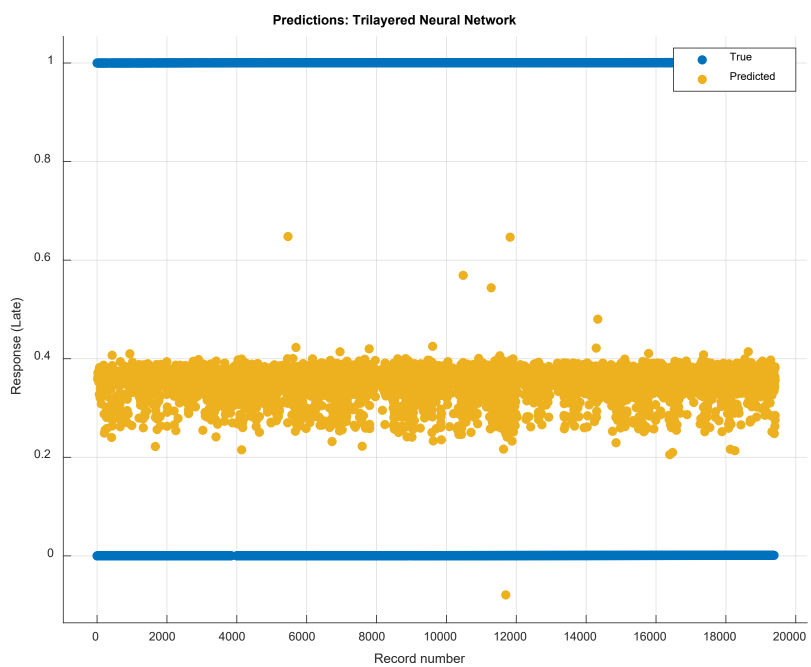

Figure 33. Graphical Representation of Trilayered Neural Network Model

The Models above were trained using all five parameters: initial cost, month of application, responsible component, CONUS/OCONUS, and associated branch. The trained models without including the associated branch, responsible component, and both the associated branch and responsible component parameters resulted in similar graphs.

C. MODEL TESTING

After training the models discussed above, we applied the models to FY21. Figures 34–38 are the graphical representations of the test results for the models including all five parameters. All the tested models showed similar results without any projects having a 0% chance of being awarded late, and very few projects having a 100% chance of being awarded late.

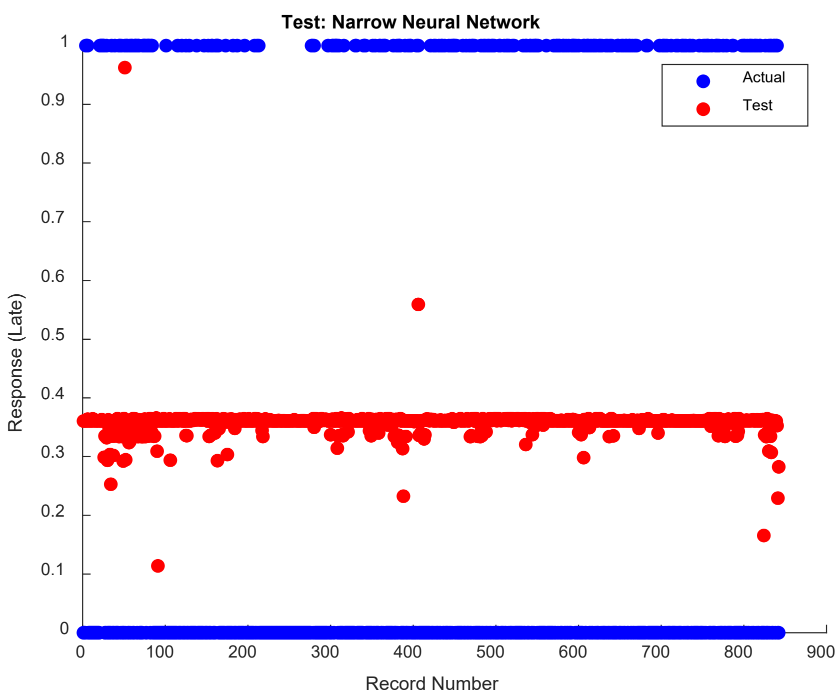

Figure 34. Test Results for Narrow Neural Network Model

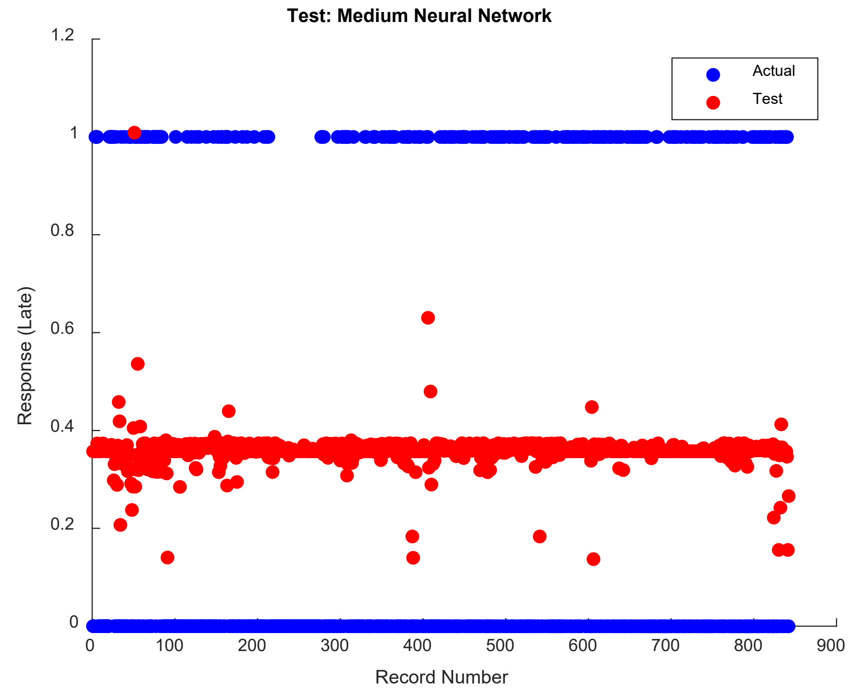

Figure 35. Test Results for Medium Neural Network Model

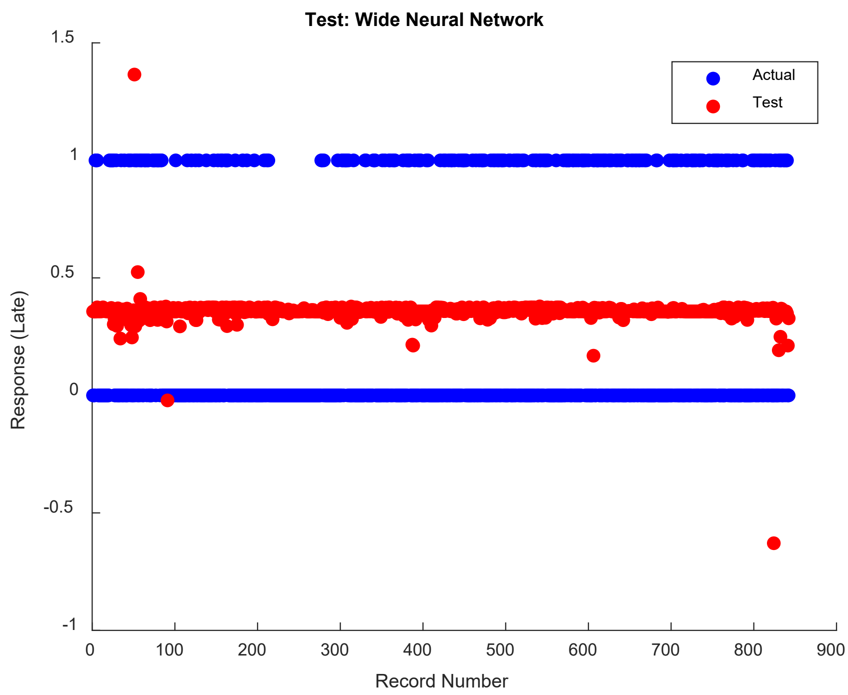

Figure 36. Test Results for Wide Neural Network Model

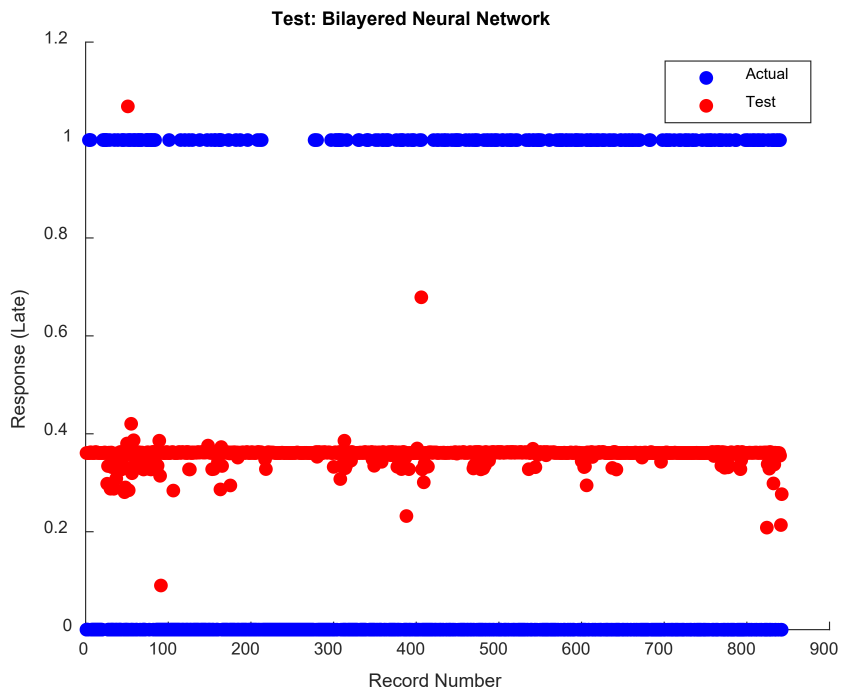

Figure 37. Test Results for Bilayered Neural Network Model

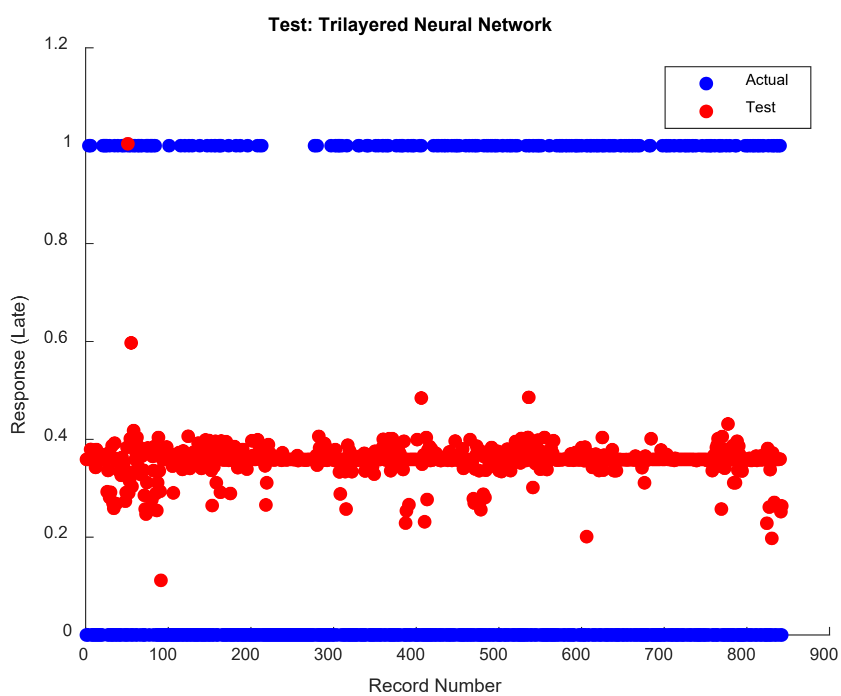

Figure 38. Test Results for Trilayered Neural Network Model

Each of the resulting figures from the tests were extremely similar to those generated from the training dataset. Looking at the test result figures above, there are very few projects with greater than a 50% chance of being awarded late. Therefore, from visual analysis of the models, the neural network models will be less accurate than the decision tree models.

D. MODEL ACCURACY

To determine the overall model accuracy, we performed the same data manipulation as for the decision tree models. We adjusted the data such that the results with values greater than 0.5 would be set to one, and the values less than 0.5 would be set to zero. The results showed that there was a minimal difference in the accuracy of each of the models, shown in Table 2. All the created neural network models had an accuracy between 72–73% with the Trilayered Neural Network model without the inclusion of the associated branch and responsible component parameters having the highest accuracy of 72.92%. Similar to the Boosted Tree model, the neural networks only predict a few projects to be late resulting the accuracy around 72% (approximately 27% of projects are late).

Table 2. Neural Network Model Test Accuracy (FY21)

ALL PARAMATERS	ACCURACY	NO BRANCH	ACCURACY	NO RESPONSIBLE COMPONENT	ACCURACY	NO BRANCH OR RESPONSIBLE COMPONENT	ACCURACY
Narrow	72.68%	Narrow	72.80%	Narrow	72.80%	Narrow	72.68%
Medium	72.80%	Medium	72.80%	Medium	72.68%	Medium	72.68%
Wide	72.68%	Wide	72.68%	Wide	72.80%	Wide	72.80%
Bilayered	72.68%	Bilayered	72.68%	Bilayered	72.80%	Bilayered	72.80%
Trilayered	72.68%	Trilayered	72.80%	Trilayered	72.68%	Trilayered	72.92%

The poor results for the neural networks, compared to the results of the decision trees, was likely due to the limitations of MATLAB’s machine learning toolbox. We were forced to use shallow neural networks (limited to three layers maximum) without the ability to adjust the weights between layers. The other likely reason for the neural network performance is the amount of data input into the training algorithm. Neural networks perform better with more data used for training. Therefore, if the entire eProjects dataset was complete, the neural networks would have been able to train 50000 projects vice 20000, thereby increasing the overall performance. Neural Networks also perform better for nontabular data; therefore decision trees are a better option for analyzing this dataset.

The confusion matrix, Figure 39, shows that the Trilayered Neural Network model provides solely true positives and false negatives. This confirms that the neural networks predominantly do not predict projects as being late for awarding.

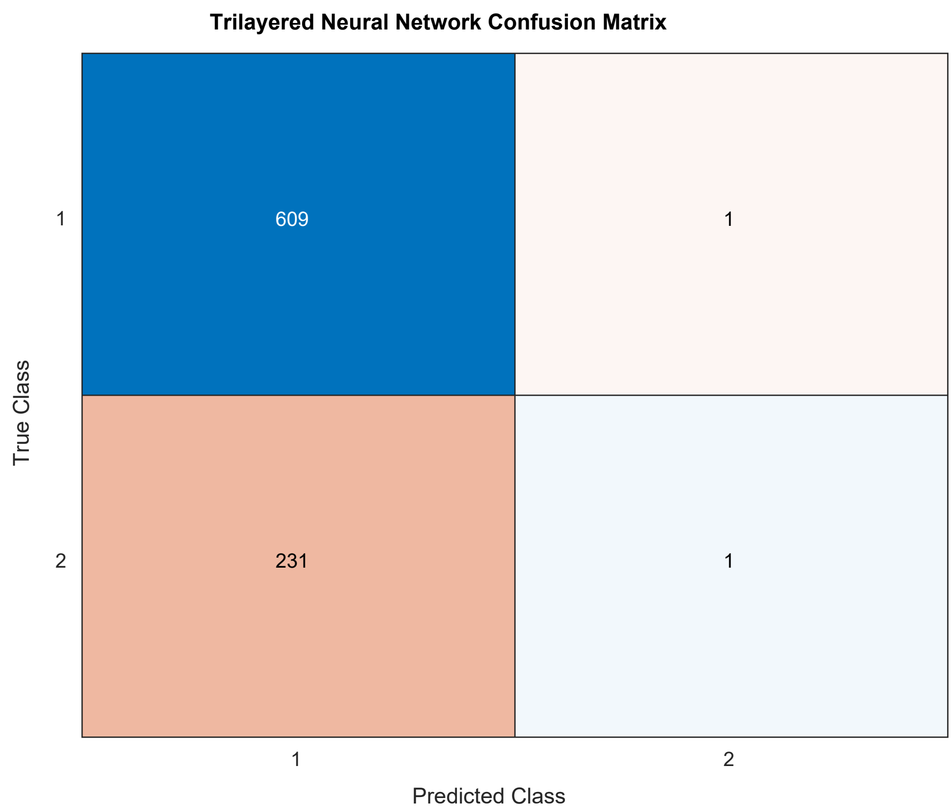

Figure 39. Confusion Matrix for Trilayered Neural Network Model Results

THIS PAGE INTENTIONALLY LEFT BLANK

VI. CONCLUSION AND FUTURE WORK

This chapter presents a summary of the conclusions derived from our statistical analysis, decision trees, and neural networks. We present the results of our findings and the possible improvements that could be made to the neural networks. The chapter ends with the possibilities for future work for the analysis and improvement of NAVFAC project timelines.

A. CONCLUSION

Overall, eProjects has an extensive amount of data. While we analyzed the data using the five most prominent parameters (initial cost, month of submission, responsible component, CONUS/OCONUS, and branch association), there were many other parameters that could have been used for analysis. The issue with using more parameters would be shrinking the overall dataset. All the data in the eProjects database must be received by NAVFAC, therefore incomplete data sent to NAVFAC limited the number of projects that we could analyze. Through the statistical analysis and machine learning algorithms we determined that the incomplete data played a large factor in the accuracy of the statistical analysis and neural network models.

Through statistical analysis, we were able to determine that there is no single parameter that causes projects to be awarded late. All the analyzed parameters resulted in a 30–40% of the project being awarded late except for the month of project submission. Therefore, statistically, the responsible components will have a higher chance of projects being awarded on time if they submit the projects in August or September. On another note, if the projects are associated with branches other than the Navy and Marine Corps there is greater than a 50% chance of projects being awarded late.

The machine learning techniques we used to create the predictive models were decision trees and neural networks. As shown in Table 3, the decision trees proved to create more accurate models than the neural networks.

Table 3. Model Accuracy for All Generated Models (Decision Trees and Neural Networks)

ALL PARAMATERS	ACCURACY	NO BRANCH	ACCURACY	NO RESPONSIBLE COMPONENT	ACCURACY	NO BRANCH OR RESPONSIBLE COMPONENT	ACCURACY
Narrow NN	72.68%	Narrow NN	72.80%	Narrow NN	72.80%	Narrow NN	72.68%
Medium NN	72.80%	Medium NN	72.80%	Medium NN	72.68%	Medium NN	72.68%
Wide NN	72.68%	Wide NN	72.68%	Wide NN	72.80%	Wide NN	72.80%
Bilayered NN	72.68%	Bilayered NN	72.68%	Bilayered NN	72.80%	Bilayered NN	72.80%
Trilayered NN	72.68%	Trilayered NN	72.80%	Trilayered NN	72.68%	Trilayered NN	72.92%
Fine Tree	94.54%	Fine Tree	95.37%	Fine Tree	95.49%	Fine Tree	92.76%
Medium Tree	88.24%	Medium Tree	88.48%	Medium Tree	90.02%	Medium Tree	87.89%
Coarse Tree	84.09%	Coarse Tree	83.37%	Coarse Tree	83.13%	Coarse Tree	80.76%
Bagged Trees	86.10%	Bagged Trees	85.75%	Bagged Trees	89.19%	Bagged Trees	86.94%
Boosted Tree	72.68%	Boosted Tree	72.92%	Boosted Tree	72.68%	Boosted Tree	72.57%

The inaccuracies in the neural networks have three likely causes. The first likely cause is that we only used shallow neural networks for our model generation. The regression learning application in the machine learning toolbox limits the neural networks to a three-layer maximum. Deep neural networks may have been able to generate more accurate models due to the parameters and dataset. If any of the analyzed parameters had been a prominent cause for delays in project award, then the neural networks would have likely been able to generate more accurate models. The next cause for the inaccuracies of the neural networks is the size of the dataset. Neural Networks work better with larger datasets. Through FY11-20 there was initially 53000 projects for analysis, but due to missing parameters across the projects we ended up with approximately 22000 projects for analysis. If we had been able to use the full 53000 projects, the neural networks would have likely been able to produce more accurate models.

In conclusion, the Fine Tree model had a 95% accuracy for predicting whether projects would be awarded on time. Follow-on analysis would be necessary to accurately apply neural networks to the data for predictive models. The only parameter that had a decisive indication for impacting the timeliness of project awarding is the month of submission. Overall, the decision tree models performed as expected as decision trees generally perform well for tabular data. The complete dataset could have provided different results as 60% of the data was incomplete. While there were few USW related projects in the eProjects database almost all of them had incomplete data. Therefore, we were unable to analyze the causes for delays in USW related project awarding.

B. FUTURE RESEARCH

Our research primarily focused on statistical analysis, decision trees, and shallow neural networks. However, we recommend applying deep neural networks to the data to determine the impact on model accuracy. Neural networks are more versatile than decision trees and would be better for long term predictive models. Aside from deep neural networks, combining the eProjects data with EPG database could produce a more complete dataset, thereby improving the accuracy of the statistical analysis and possibly the neural networks.

We also recommend further analysis into the NAVFAC databases. NAVFAC is interested in the causes for overall project delays. While the eProjects database only covers the projects until project award, the Electronic Construction and Facility Support Contracts Management System (eCMS) database follows the projects from award to completion. Therefore, we recommend further research and analysis into the eCMS database to determine likely causes for project delays and for correlations between project parameters and delays in project completion.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Brown K, Dagan H, Kidda B, Ogata D (2020) Step 3: Initiate project. Lecture MILCON Installation/PWD1391, July 14, Naval Facilities Engineering Command Process Driven Training, https://totalforcetraining.navfac.navy.mil/

Department of Navy (2018), A design for maintaining maritime superiority. Strategic Design ver 2.0, Washington, DC. https//www.navy.mil/ah_online/ MaritimeSuperiority/

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.

Iefacman. (2022). Naval Facilities Engineering Command eProjects Database. Retrieved March 18, 2022, from https://iefacman.navfac.navy.mil/.

Kamiński, B., Jakubczyk, M., & Szufel, P. (2017). A framework for sensitivity analysis of Decision Trees. *Central European Journal of Operations Research*, 26(1), 135–159. https://doi.org/10.1007/s10100-017-0479-6

Kurama, V. (2021, April 9). *A guide to understanding AdaBoost*. Paperspace Blog. Retrieved July 29, 2022, from https://blog.paperspace.com/adaboost-optimizer/#:~:text=AdaBoost%20is%

Komiss, W. C., and Saulo, R. L. C. (2021). NAVFAC Performance to Plan for Military Construction. CNA.

Mathworks. (n.d.). *Choose Regression Model Options*. Choose Regression Model Options - MATLAB and Simulink. Retrieved August 5, 2022, from https://www.mathworks.com/help/stats/choose-regression-model-options.html#mw_1d867187-1f2b-4711-bde6-75d18d9621be

Mathworks (n.d.). *Decision Trees*. Choose Classifier Options - MATLAB and Simulink. Retrieved July 29, 2022, from https://www.mathworks.com/help/stats/choose-a-classifier.html

Mathworks Inc. (n.d.). *Regression Learner Application*. Train Regression Models in Regression Learner App - MATLAB & Simulink. Retrieved August 23, 2022, from https://www.mathworks.com/help/stats/train-regression-models-in-regression-learner-app.html

Naval Facilities Engineering Command (2019) External summary memo. *Strategic Design 2.0*. Washington, DC. https://hub.navfac.navy.mil/webcenter-/portal/ Strategic_Design

Thompson, R. J. (2020). *High-Influence Factors for the Timeliness of Project Award for Navy Military Construction* (thesis). Naval Postgraduate School, Monterey.

THIS PAGE IS INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California