



NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

EMPLOYING MACHINE LEARNING TO PREDICT STUDENT

AVIATOR PERFORMANCE

by

Dr. Magdi N Kamel

October 2020

Approved for public release; distribution is unlimited

Prepared for: Naval Research Program

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 instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this 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 Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE (DD-MM-YYYY) 14-10-2020		2. REPORT TYPE Final		3. DATES COVERED (From-To) Oct 2019 – Oct 2020	
4. TITLE AND SUBTITLE Employing Machine Learning to Predict Student Aviator Performance				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Kamel, Magdi, N				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) AND ADDRESS(ES) Naval Postgraduate School 1 University Circle Monterey, CA 93943				8. PERFORMING ORGANIZATION REPORT NUMBER NPS-IS-20-002	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Sponsor: Naval Research Program Topic Sponsor: Commander, U.S. Pacific Fleet (COMPACFLT) Sponsor Address: Naval Postgraduate School, 1 University Circle, Monterey, CA 93943 Project JON: W0A34 NRP Project (IREF) ID: NPS-20-N309-A				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT Machine learning analysis of student aviator training performance data offers novel and more accurate methodologies for performance assessment that includes identifying students for attrition or remediation as well as optimal pipeline assignments. Machine learning provides an opportunity to better evaluate students by fully examining every indicator of performance throughout a student's training: from subtest scores on the aviation selection test battery to test scores during initial ground school through each graded item on every flight event. The goal is to reduce time-to-train, improve aviator quality, and reduce training costs from failure to complete training. In this research, we identify important predictors and develop prediction models of performance in Primary, Intermediate, and Advanced flight training based on data from Aviation Selection Test Battery (ASTB), Introductory Flight Screening (IFS), and Aviation Preflight Indoctrination (API) training.					
15. SUBJECT TERMS Machine learning, Data analytics, Predictive models, Aviation training					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 101	19a. NAME OF RESPONSIBLE PERSON Kamel, Magdi, N
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U			19b. TELEPHONE NUMBER (include area code) 831-656-2494

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

THIS PAGE INTENTIONALLY LEFT BLANK

**NAVAL POSTGRADUATE SCHOOL
Monterey, California 93943-5000**

Ann E. Rondeau
President

Rob Dell
Provost

The report entitled “Employing Machine Learning to Predict Student Aviator Performance” was prepared for Navy Research Program and funded by Navy Research Program.

Further distribution of all or part of this report is authorized.

or

Further distribution of all or part of this report is subject to the Distribution Statement appearing on the front cover.

This report was prepared by:

Dr. Magdi N Kamel
Associate Professor

Author
Title

Reviewed by:

Released by:

Alex Bordetsky, Chairman
Information Sciences

Jeffrey D. Paduan
Dean of Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Machine learning analysis of student aviator training performance data offers novel and more accurate methodologies for performance assessment that includes identifying students for attrition or remediation as well as optimal pipeline assignments. Machine learning provides an opportunity to better evaluate students by fully examining every indicator of performance throughout a student's training: from subtest scores on the aviation selection test battery to test scores during initial ground school through each graded item on every flight event. The goal is to reduce time-to-train, improve aviator quality, and reduce training costs from failure to complete training. In this research, we identify important predictors and develop prediction models of performance in Primary, Intermediate, and Advanced flight training based on data from Aviation Selection Test Battery (ASTB), Introductory Flight Screening (IFS), and Aviation Preflight Indoctrination (API) training.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

I. INTRODUCTION.....	17
A. BACKGROUND	17
B. NAVAL AVIATION ORDER OF TRAINING	17
1. Aviation Selection Test Battery (ASTB)	18
2. Introductory Flight Screening (IFS)	18
3. Aviation Preflight Indoctrination (API)	18
4. Primary Flight Training (PRI)	18
5. Intermediate Flight Training (INT)	18
6. Advanced Flight Training (ADV).....	19
7. Fleet Replacement Squadron (FRS).....	19
C. RESEARCH/ STUDY AND ANALYSIS OBJECTIVES	20
D. DATA SCIENCE METHODOLOGY	20
1. Business understanding.....	21
2. Data understanding	21
3. Data preparation.....	21
4. Modeling	21
5. Evaluation.....	21
6. Deployment.....	22
II. DATA UNDERSTANDING AND PREPARATION.....	23
A. DATA UNDERSTANDING	23
B. DATA PREPARATION	23
1. ASTB345 Data Sets	23
a. ASTB345 Primary Data Set.....	23
b. ASTB345 Intermediate Data Set	23
c. ASTB345 Advanced Data Set	24
2. ASTBE Data Sets	24
a. ASTBE Primary Data Set.....	24
a. ASTBE Intermediate Data Set.....	24
b. ASTBE Advanced Data Set.....	24
3. Missing Values.....	24
4. Data Preparation Tasks.....	25
III. MODELING ASTB345 DATA TO PREDICT PRIMARY TRAINING	
OUTCOME	26
1. Predicting Outcome of Primary Training (PRI) from ASTB345 Scores... 26	
a. Multiple Logistics Model with Backward Elimination.....	27
b. Bootstrap (Random) Forest Model.....	30
c. Penalized Regression Models	32
d. Model Comparison.....	39
e. Conclusions	40
2. Predicting Outcome of Primary Training (PRI) from IFS Scores..... 41	
a. Multiple Logistics Model with Backward Elimination.....	42

b.	<i>Bootstrap (Random) Forest Model</i>	44
c.	<i>Penalized Regression Model</i>	45
d.	<i>Model Comparison</i>	48
e.	<i>Conclusions</i>	48
3.	Predicting Outcome of Primary Training (PRI) from API Scores	49
a.	<i>Multiple Logistics Model</i>	49
b.	<i>Decision (Partition) Tree</i>	51
c.	<i>Model Comparison</i>	53
d.	<i>Conclusions</i>	54
IV.	MODELING ASTBE DATA TO PREDICT PRIMARY TRAINING OUTCOME	55
1.	Predicting Outcome of Primary Training (PRI) from ASTBE Scores	55
a.	<i>Multiple Logistics Model with Backward Elimination</i>	56
b.	<i>Bootstrap (Random) Forest Model</i>	58
c.	<i>Model Comparison</i>	62
d.	<i>Conclusions</i>	62
2.	Predicting Outcome of Primary Training (PRI) from IFS Scores	63
a.	<i>Multiple Logistics Model with Backward Elimination</i>	63
b.	<i>Bootstrap (Random) Forest Model</i>	65
c.	<i>Model Comparison</i>	67
d.	<i>Conclusions</i>	67
3.	Predicting Outcome of Primary Training (PRI) from API Scores	68
a.	<i>Multiple Logistics Model</i>	68
b.	<i>Decision (Partition) Tree</i>	70
c.	<i>Model Comparison</i>	71
d.	<i>Conclusions</i>	72
V.	MODELING ASTBE DATA TO PREDICT ADVANCED TRAINING OUTCOME	73
1.	Predicting Outcome of Primary Training (ADV) from ASTBE Scores	73
a.	<i>Multiple Logistics Model with Backward Elimination</i>	73
b.	<i>Bootstrap (Random) Forest Model</i>	75
c.	<i>Model Comparison</i>	79
d.	<i>Conclusions</i>	79
2.	Predicting Outcome of Advanced Training (ADV) from IFS Scores	80
a.	<i>Multiple Logistics Model with Backward Elimination</i>	80
b.	<i>Bootstrap (Random) Forest Model</i>	81
c.	<i>Model Comparison</i>	83
d.	<i>Conclusions</i>	84
3.	Predicting Outcome of Advanced Training (ADV) from API Scores	84
a.	<i>Multiple Logistics Model</i>	84
b.	<i>Decision (Partition) Tree Model</i>	86
c.	<i>Model Comparison</i>	88
d.	<i>Conclusions</i>	89
VI.	SUMMARY AND CONCLUSIONS	91
A.	SUMMARY	91

1.	Predicting outcome of PRI training using ASTB345 Data.....	91
a.	From ASTB345 Scores	91
b.	From IFS Scores.....	91
c.	From API Scores.....	92
2.	Predicting outcome of PRI training using ASTBE Data.....	92
a.	From ASTBE Scores.....	92
b.	From IFS Scores.....	92
c.	API Scores.....	93
3.	Predicting outcome of ADV training using ASTBE Data	93
a.	From ASTBE Scores.....	93
b.	From IFS Scores.....	93
c.	From API Scores.....	94
B.	CONCLUSIONS	94
	LIST OF REFERENCES	95
	INITIAL DISTRIBUTION LIST	97

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

Figure 1. Naval aviation order of training	17
Figure 2. The data mining life cycle	20
Figure 3. Correlations of ASTB345 predictors.....	27
Figure 4. Logistics model effect summary for ASTB345.....	27
Figure 5. Logistics model whole model test report for ASTB345.....	28
Figure 6. Logistics model fit details report for ASTB345	28
Figure 7. Logistics model confusion matrix for ASTB345	28
Figure 8. Logistics model decision matrix for ASTB345.....	29
Figure 9. Logistics model prediction profiler for ASTB345	30
Figure 10. Bootstrap forest overall model summary statistics for ASTB345.....	30
Figure 11. Bootstrap forest model column contribution for ASTB345.....	31
Figure 12. Bootstrap forest model confusion matrix for ASTB345	31
Figure 13. Bootstrap forest model decision Matrix for ASTB345	31
Figure 14. Bootstrap forest model prediction profiler for ASTB345	32
Figure 15. Lasso model for ASTB345.....	34
Figure 16. Lasso model prediction profiler for ASTB345.....	34
Figure 17. Elastic net model for ASTB345	36
Figure 18. Elastic net model prediction profiler for ASTB345	37
Figure 19. Ridge model for ASTB345.....	39
Figure 20. Ridge model prediction profiler for ASTB345	39
Figure 21. Model comparison on training data set for ASTB345.....	40
Figure 22. Model comparison on validation data set for ASTB345.....	40
Figure 23. Correlation of IFS scores.....	41
Figure 24. Logistics model effect summary for IFS	42
Figure 25. Logistics model whole model test	42
Figure 26. Logistics model confusion matrix for IFS.....	43
Figure 27. Logistics model decision matrix for IFS	43
Figure 28. Logistics model prediction profiler for IFS.....	44
Figure 29. Bootstrap forest overall model summary for IFS	44
Figure 30. Bootstrap forest model column contribution for IFS	44
Figure 31. Bootstrap forest model confusion matrix for IFS.....	45
Figure 32. Bootstrap forest model prediction profiler for IFS.....	45
Figure 33. Elastic net model for IFS	47
Figure 34. Elastic net model prediction profiler for IFS.....	48
Figure 35. Model comparison for training set for IFS	48
Figure 36. Model comparison for validation set for IFS	48
Figure 37. Logistics model effect summary for API	49
Figure 38. Logistics model whole model test for API.....	50
Figure 39. Logistics model confusion matrix for API	50
Figure 40. Logistics model prediction profiler for API	51
Figure 41. Decision tree model for API.....	52
Figure 42. Decision tree model column contribution for API	53

Figure 43. Decision tree model fit details and confusion matrices for API.....	53
Figure 44. Decision tree model leaf report for API	53
Figure 45. Model comparison for training set for API	54
Figure 46. Model comparison for validation set for API.....	54
Figure 47. Correlations of ASTBE predictors	56
Figure 48. Logistics model effect summary for ASTBE	57
Figure 49. Logistics model whole model test report for ASTBE	57
Figure 50. Logistics model fit details report for ASTBE.....	57
Figure 51. Logistics model confusion matrix for ASTBE.....	58
Figure 52. Logistics model prediction profiler for ASTBE.....	58
Figure 53. Bootstrap forest overall model summary statistics for ASTBE	59
Figure 54. Bootstrap forest model column contributions for ASTBE.....	59
Figure 55. Bootstrap forest model confusion matrix for ASTBE.....	60
Figure 56. Bootstrap forest model prediction profiler for ASTBE.....	61
Figure 57. Model comparison for training set for ASTBE	62
Figure 58. Model comparison for validation set for ASTBE.....	62
Figure 59. Correlation of IFS scores.....	63
Figure 60. Logistics model effect summary for IFS	64
Figure 61. Logistics model whole model test for IFS.....	64
Figure 62. Logistics model confusion matrix for IFS.....	64
Figure 63. Logistics model prediction profiler for IFS.....	65
Figure 64. Bootstrap forest overall model summary for IFS	65
Figure 65. Bootstrap forest model column contributions for IFS	66
Figure 66. Bootstrap forest model confusion matrix for IFS.....	66
Figure 67. Bootstrap forest model prediction profiler for IFS.....	67
Figure 68. Model comparison for training set for IFS	67
Figure 69. Model comparison for validation set for IFS	67
Figure 70. Logistics model effect summary for API	68
Figure 71. Logistics model whole model test for API.....	69
Figure 72. Logistics model confusion matrix for API	69
Figure 73. Logistics model prediction profiler for API	70
Figure 74. Decision tree for API.....	70
Figure 75. Decision tree model column contribution for API	71
Figure 76. Decision tree model fit details and confusion matrices for API.....	71
Figure 77. Decision tree model leaf report for API	71
Figure 78. Model comparison for training set for API	72
Figure 79. Model comparison for validation set for API.....	72
Figure 80. Logistics model effect summary for ASTBE	73
Figure 81. Logistics model whole model test report for ASTBE	74
Figure 82. Logistics model fit details report for ASTBE.....	74
Figure 83. Logistics model confusion matrix for ASTBE.....	74
Figure 84. Logistics model prediction profiler for ASTBE.....	75
Figure 85. Bootstrap forest overall model summary statistics for ASTBE	75
Figure 86. Bootstrap forest model column contribution for ASTBE.....	76
Figure 87. Bootstrap forest model confusion matrix for ASTBE.....	76
Figure 88. Bootstrap forest model prediction profiler for ASTBE.....	78

Figure 89. Model comparison for training set for ASTBE	79
Figure 90. Model comparison for validation set for ASTBE.....	79
Figure 91: Logistics model effect summary for IFS	80
Figure 92. Logistics model whole model test for IFS.....	80
Figure 93. Logistics model confusion matrix for IFS.....	81
Figure 94. Logistics model prediction profiler for IFS.....	81
Figure 95. Bootstrap forest overall model summary for IFS	82
Figure 96. Bootstrap forest model column contribution for IFS	82
Figure 97. Bootstrap forest model confusion matrix for IFS.....	82
Figure 98. Bootstrap forest model prediction profiler for IFS.....	83
Figure 99. Model comparison for training set for IFS	83
Figure 100. Model comparison for validation set for IFS	83
Figure 101. Logistics model effect summary for API	84
Figure 102. Logistics model whole model test for API.....	85
Figure 103. Logistics model confusion matrix for API.....	85
Figure 104. Logistics model prediction profiler for API	86
Figure 105. Decision tree for API.....	87
Figure 106. Decision tree model column contribution for API	88
Figure 107. Decision tree model fit details and confusion matrices for API.....	88
Figure 108. Decision tree model leaf report for API	88
Figure 109. Model comparison for training set for API	89
Figure 110. Model comparison for validation set for API.....	89

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

Table 1. Training phases predictor and target variables	19
---	----

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. BACKGROUND

For decades, naval aviation training has used the Navy Standard Score (NSS) as the primary means of overall performance evaluation of its students. With the current shift into new training management systems (i.e., Training Sierra Hotel Aviation Readiness Program [T-SHARP]), new methodologies for evaluating and predicting student performance should be examined. Machine learning provides an opportunity to better evaluate students by fully examining every indicator of performance throughout a student's training, from subtest scores on the aviation selection test battery and test scores during initial ground school through each graded item on every flight event. This full integration of performance criteria will identify trends and patterns traditional methods lack, providing a better overall evaluation of whether a student will be successful in training, which aviation pipeline a student will be most successful, and which students need remediation earlier, in order to provide additional resources for success.

B. NAVAL AVIATION ORDER OF TRAINING

The following is a brief description of the order of training for Naval Aviation Flight training (Natali, 2019) as depicted in Figure 1 (Chief of Naval Air Training, n.d.).



Figure 1. Naval aviation order of training

1. Aviation Selection Test Battery (ASTB)

The most recent iteration of the test, ASTB-E, went operational in December 2013. This is the primary selection assessment of aviation candidates and is taken before entering aviation training. Candidates must achieve minimum scores to be considered for further training. It consists of assessments of cognitive ability, personality, and psychomotor skills. A detailed description of the components of the test can be found in (Navy Medicine Operational Training Center (NMOTC), n.d.).

2. Introductory Flight Screening (IFS)

This initial aviation training in a civilian aircraft is intended to give students a feel for how aviation is. It consists of academics with multiple tests and approximately 13-15 hours of flight time.

3. Aviation Preflight Indoctrination (API)

The API consists of academics on basic aviation concepts with tests on the following: Aerodynamics (two parts), Weather, Engine Systems, Flight Rules and Regulations, and Navigation. Candidates must achieve an 80 (0-100 scale) to pass a given test, but must achieve a final Navy Standardized Score (NSS) above 35 (20-80 scale) to complete. Since it is standardized among peers, average test scores are in the 90s to remain competitive and not fall below the 35 NSS threshold. Additionally, within NSS, the six tests are weighted, with Aerodynamics and Navigation having greater weights than the other three.

4. Primary Flight Training (PRI)

This is the first phase of military Navy flight training with the T-6B Texan aircraft, and consists of graded academics, simulator events, and flight events. Overall performance of students is reported via NSS and is standardized across the most recent 200 completers. Performance in Primary is the main determinant of pipeline/aircraft assignment: Jet-Strike, Jet-E2/C2, Rotary, Multi-Engine, and Tiltrotor.

5. Intermediate Flight Training (INT)

This training is the second phase of military Navy flight training, and is only for specific pipelines: Strike, E2/C2, and Tiltrotor. It consists of graded academics, simulator

events, and flight events. Overall performance of students is reported via NSS and standardized across the most recent 60 completers.

6. Advanced Flight Training (ADV)

This is the final phase of flight training before being assigned operational aircraft. Advanced training is completed in either the T-45C (jet) for Strike, T-44 (multi-engine) for Multi-engine, E2/C2, and Tilt, or TH-57 (rotary) for helicopters. It consists of graded academics, simulator events, and flight events. Overall performance of students is reported via NSS and standardized across the most recent 60 completers.

7. Fleet Replacement Squadron (FRS)

This is the operational aircraft training on newly winged aviators (i.e., graduated CNATRA training). This training covers several squadrons and several types of aircraft across pipelines.

Only two datasets are either only an Independent (predictor) Variable (IV) or Dependent (target) Variable (DV): ASTB scores are only an IV and FRS data only a DV. All data in between these two can be both predictor and criterion dependent on what training has been completed, i.e., nested within each other (Table 1).

Table 1. Training phases predictor and target variables

	ASTB	IFS	API	PRI	INT	ADV
IFS	X					
API	X	X				
PRI	X	X	X			
INT	X	X	X	X		
ADV	X	X	X	X	X	
FRS	X	X	X	X	X	X

C. RESEARCH/ STUDY AND ANALYSIS OBJECTIVES

The long-term objective of this research is to conduct analytics on collected naval aviation training data using statistical and machine learning techniques, in order to answer the following research questions:

1. What correlations exist between the collected performance indicators and performance during assessments?
2. What performance indicators differentiate student performance between training pipelines (SNA vs. SNFO; Jet vs. Rotary vs. Multi-Engine)?
3. Can predictive models of aviator performance be developed based on the identified performance indicators?
4. What is the earliest point where analytics differentiate which aircraft pipeline(s) students will be most successful?

In this effort, we identify important predictors and develop prediction models of performance in Primary, Intermediate, and Advanced flight training based on data from Aviation Selection Test Battery (ASTB), Introductory Flight Screening (IFS), and Aviation Preflight Indoctrination (API) training.

D. DATA SCIENCE METHODOLOGY

The data science methodology used to guide the analytics of this research is based on the Cross-Industry Standard Process for Data Mining (CRISP DM). Figure 2 depicts the life cycle of the methodology (Vorhies, 2016).

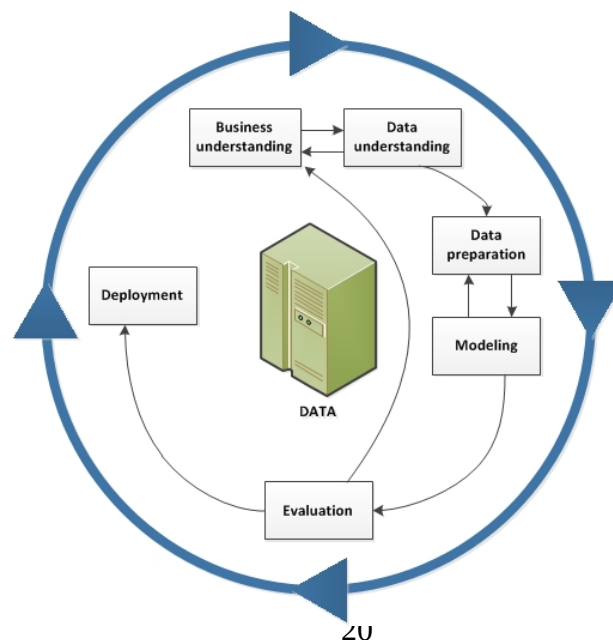


Figure 2. The data mining life cycle

The life cycle model consists of six phases: Business understanding, Data understanding, Data preparation, Modeling, Evaluation, and Deployment with arrows indicating the most important and frequent dependencies between phases. The execution of the phases is not strictly sequential; most projects move back and forth between phases as necessary.

1. Business understanding

This is perhaps the most important phase of data mining. Business understanding includes determining business objectives, assessing the situation, determining data mining goals to support business objectives, and producing a project plan.

2. Data understanding

This phase addresses the need to understand available data resources and the characteristics of those resources. It includes collecting initial data, describing data, exploring data, and verifying data quality.

3. Data preparation

After cataloging data resources, data need to be prepared for mining. Preparations include selecting, cleaning, constructing, integrating, and formatting data. These tasks are likely performed multiple times, and not in any prescribed order. These tasks can be very time consuming but are critical for the success of the data-mining project.

4. Modeling

This is the most interesting part of data mining, where sophisticated analysis methods are used to extract information from the data. This phase involves selecting modeling techniques, generating test designs, and building and assessing models.

5. Evaluation

Before writing final reports and deploying the model, it is important to thoroughly evaluate the model, and review the steps executed to construct the model, to be certain it properly achieves the business objectives. A key aim is to determine if there is some important business issue that has not been sufficiently considered. At the end of this phase, a decision will be made on the use of the data-mining results.

6. Deployment

The deployment phase can be as simple as generating a report or as complex as implementing a repeatable data-mining process. However, the focus of this effort is on the data preparation, data understanding, and modeling phases of the methodology.

II. DATA UNDERSTANDING AND PREPARATION

A. DATA UNDERSTANDING

The ASTB data set received consisted of 14,901 observations with 116 variables with data from ASTB12, ASTB345, and ASTBE all in one file. Each observation represents a student aviator scores on ASTB (12, 345, or E), IFS, API and the graduation status for Primary (PRI), Intermediate (INT), and Advanced (ADV) flight training. Graduation status include: Graduate (G), Under Instruction (UI), Not there yet (AT), Not Applicable (NA), Non-Graduate/Attrite (NG).

B. DATA PREPARATION

The first data preparation task undertaken was to extract ASTB12, ASTB345, and ASTBE observations into separate data sets for separate analysis since questions for each type of ASTB are different. Further, Primary, Intermediate, and Advanced data were extracted from each subset and only those observations with Graduate (G) and Non-Graduate (NG) outcomes were included in these subsets. ASTB12 data was not analyzed further since ASTB12 is no longer used and is considered obsolete. The resulting data subsets are described in the following subsections.

1. ASTB345 Data Sets

a. ASTB345 Primary Data Set

This data set consisted of 8,753 observations with 91 variables. It represented the scores of student aviators on ASTB345, IFS, API, for all student aviators with a graduation status in primary training of G or NG. Note that observations with graduation status other than G and NG were removed as they are not of interest to predict.

b. ASTB345 Intermediate Data Set

This data set consisted of 3,624 observations with 91 variables. It represented the scores of student aviators on ASTB345, IFS, API, for all student aviators with graduation status in intermediate training of G or NG. Similarly, observations with graduation status other than G and NG were removed as they are not of interest to predict.

c. *ASTB345 Advanced Data Set*

This data set consisted of 5,387 observations with 91 variables. It represented the scores of student aviators on ASTB345, IFS, API, for all student aviators with graduation status in advanced training of G or NG. Observations with graduation status other than G and NG were removed as they are not of interest to predict.

2. *ASTBE Data Sets*

a. *ASTBE Primary Data Set*

This data set consisted of 2,311 observations with 98 variables. It represented the scores of student aviators on ASTBE, IFS, API, for all student aviators with graduation status in primary training of G or NG. Note that observations with graduation status other than G and NG were removed as they are not of interest to predict.

a. *ASTBE Intermediate Data Set*

This data set consisted of 682 observations with 98 variables. It represented the scores of student aviators on ASTBE, IFS, API, for all student aviators with a graduation status in intermediate training of G or NG. Similarly, observations with graduation status other than G and NG were removed as they are not of interest to predict.

b. *ASTBE Advanced Data Set*

This data set consisted of 1,181 observations with 98 variables. It represented the scores of student aviators on ASTBE, IFS, API, for all student aviators with graduation status in advanced training of G or NG. Observations with graduation status other than G and NG were removed as they are not of interest to predict.

3. *Missing Values*

Missing value patterns were identified in the data sets, and particularly evident in IFS scores data. Based on these patterns, we elected to use informative missing, a coding system, to address missing values. The system allows estimation of a predictive model despite the presence of missing values. For linear models, when informative missing is used the missing values are handled as follows (SAS Institute Inc., 2018, pp 55–56):

- When a categorical (nominal or ordinal) variable has missing values, the missing values are coded as a separate level of that variable

- When a continuous variable has missing values, they are replaced with the mean of the non-missing values for rows included in the analysis. The mean is a neutral value that maintains the interpretability of parameter estimates.

For partition models, when informative missing is used the missing values are handled as follows (SAS Institute Inc., 2018, pp 104–105):

- Rows containing missing values for a categorical (nominal or ordinal) predictor are entered into the analysis as a separate level of the variable.
- Rows containing missing values for a continuous predictor are assigned to a split as follows: The values of the continuous predictor are sorted. Missing rows are first considered to be on the low end of the sorted values. All splits are constructed. The missing rows are then considered to be on the high end of the sorted values. Again, all splits are constructed. The optimal split is determined using the LogWorth criterion. For further splits on the given predictor, the algorithm commits the missing rows to high or low values, as determined by the first split induced by that predictor.

4. Data Preparation Tasks

Subsequent data preparation tasks were undertaken to address data quality issues in the data sets, including: incorrect formatting, incomplete data, missing data, and dirty or messy data. The following is a partial list of the data preparation tasks undertaken to address the identified data quality issues:

1. Correctly specified data and modeling types for variables
2. Recoded the value “m” to “M” in Gender column
3. Recoded the value “Cauc” to “White” in Race column
4. Recoded the values “TG” and “UA” to “NG” and “UI” respectively in Pri column
5. Recoded the values “J” to “G” in Int column
6. Recoded the values “AQ”, “TG”, “UIT”, and “UU” to “AT”, “NG”, “UI” and “UI” respectively in adv column
7. Recoded the values “M” and “Uhello” to “Multi” and “Helo” respectively in Select column
8. Recoded the values “Disenroll” to “Disenroll” in IFS_STATUS column
9. Recoded unknown value “370482” to missing in the API_Service column

III. MODELING ASTB345 DATA TO PREDICT PRIMARY TRAINING OUTCOME

In this chapter we use ASTB345 data to predict outcome of Primary training. The data set for this analysis consists of 8,753 observations and 91 variables.

1. Predicting Outcome of Primary Training (PRI) from ASTB345 Scores

In this section, we develop multiple logistics models, bootstrap (random) forest model, and several penalized regression models (Lasso, Elastic Net, and Ridge) to predict the outcome of primary training (PRI) (Graduate or Non-Graduate/Attrite) from a set of predictors that include Gender, Race, Hours of previous formal flight instruction, Recruiting Branch, Examinee status, and a set of ASTB345 exam scores data. Observations including values other than Graduate (G) or Non-Graduate (NG) are removed as they are not of interest for this model. We note that some of the ASTB345 predictors, namely the z-scores for Academic Qualifications Rating (AQR), Pilot Flight Aptitude Rating (PFAR), Flight Officer Flight Aptitude Rating (FOFAR), and Officer Aptitude Rating (OAR) are highly correlated, as shown in Figure 3.

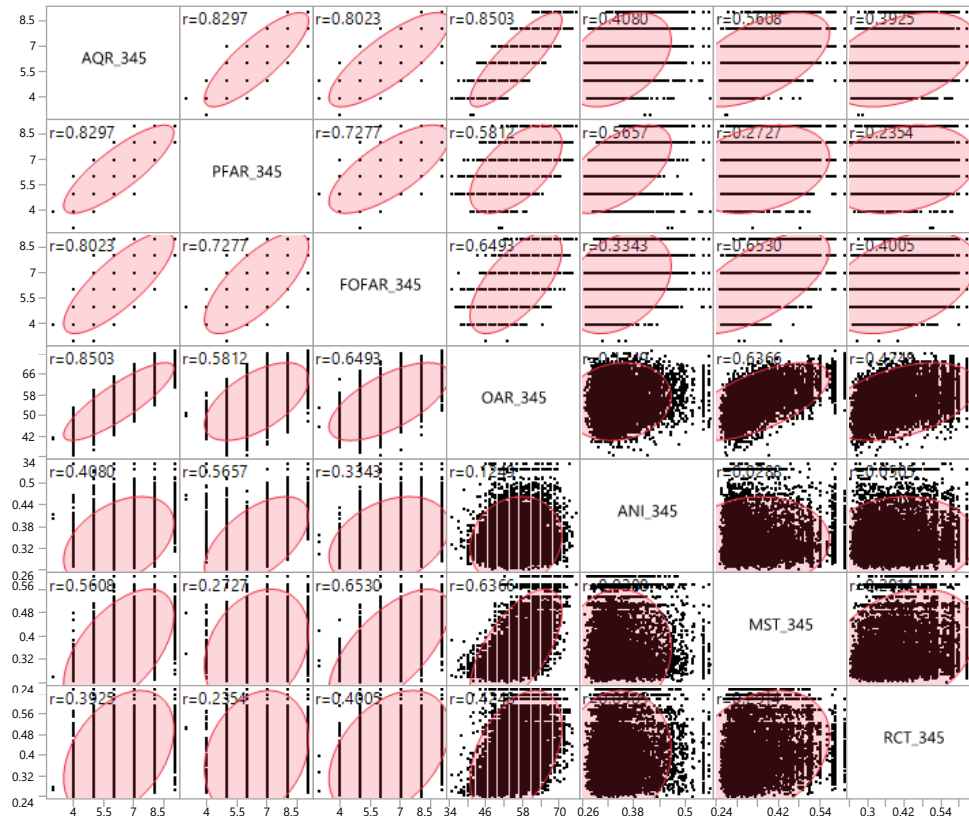


Figure 3. Correlations of ASTB345 predictors

a. Multiple Logistics Model with Backward Elimination

This multiple logistics model predicts the outcome of primary training from the full set of potential predictors (main effects). The model is then reduced using backward elimination to remove non-significant predictors. Figure 4 shows the effect summary of the model after removing the insignificant predictors.

Source	LogWorth	PValue
AQR_RAW_345	8.994	0.00000
OAR_RAW_345	6.270	0.00000
ANI_RAW_345	3.241	0.00057

Figure 4. Logistics model effect summary for ASTB345

As shown in Figure 4, the top three significant effects for predicting the outcome of primary training are the z-scores for Academic Qualifications Rating (AQR), Officer Aptitude Rating (OAR), and Aviation and Nautical Information (ANI). The

Whole Model Test Report, shown in Figure 5, has a small p-value (<0.0001), which indicates that the overall model is significant, and an R^2 of 0.0272 indicating that only 2.7% of the variation is explained by the model.

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	49.4591	3	98.91819	<.0001*
Full	1765.7229			
Reduced	1815.1820			

RSquare (U)	0.0272
AICc	3539.45
BIC	3566.33
Observations (or Sum Wgts)	6125

Figure 5. Logistics model whole model test report for ASTB345

The Fit Details Report, shown in Figure 6, indicates the model has a misclassification rate of about 8.7% on both the training and validation sets. However, this low misclassification rate is misleading due to the data imbalance in the response variable, since the number of Non-Graduates is only about 9% vs. 91% for Graduates.

Measure	Training	Validation	Definition
Entropy RSquare	0.0272	0.0195	$1 - \text{Loglike}(\text{model}) / \text{Loglike}(0)$
Generalized RSquare	0.0358	0.0257	$(1 - (L(0)/L(\text{model}))^{(2/n)}) / (1 - L(0)^{(2/n)})$
Mean -Log p	0.2883	0.2911	$\sum -\text{Log}(p[j]) / n$
RMSE	0.2800	0.2810	$\sqrt{\sum (y[j] - p[j])^2 / n}$
Mean Abs Dev	0.1568	0.1564	$\sum y[j] - p[j] / n$
Misclassification Rate	0.0873	0.0876	$\sum (p[j] \neq pMax) / n$
N	6125	2626	n

Figure 6. Logistics model fit details report for ASTB345

Figure 7 shows the confusion matrix for both the training and validation sets.

Training			Validation		
Actual	Predicted Count		Actual	Predicted Count	
Pri	NG	G	Pri	NG	G
NG	0	535	NG	0	230
G	0	5590	G	0	2396

Figure 7. Logistics model confusion matrix for ASTB345

As a result of the data imbalance, the model does not predict any Non-Graduates. To remedy this situation and specify a better classifier, a cut off lower than 0.5 needs to be specified for classifying a predicted probability for Non-Graduates. For example, the confusion matrix for a cut off of 0.1 is shown in Figure 8. The classifier of this model now predicts Non-Graduates but with a higher overall misclassification rate.

Training			Validation			Specified Profit Matrix		
Actual			Actual			Decision		
Pri	NG	G	Pri	NG	G	Actual	NG	G
NG	266	269	NG	108	122	NG	0	-1
G	1744	3846	G	711	1685	G	-0.111	0
Actual			Actual					
Pri	NG	G	Pri	NG	G			
NG	0.497	0.503	NG	0.470	0.530			
G	0.312	0.688	G	0.297	0.703			
Misclassification Rate			Misclassification Rate					
0.3287			0.3172					

Figure 8. Logistics model decision matrix for ASTB345

Figure 9 shows the prediction profiler. It indicates the predicted probability of the response variable PRI for each specified value of the predictor variables. As shown in the figure, a higher z-score for Academic Qualifications Rating (AQR) indicates a lower probability of Non-Graduate outcome, while a higher z-score for both the Officer Aptitude Rating (OAR), and Aviation and Nautical Information (ANI) is associated with a higher probability of Non-Graduate. We note however that with multicollinearity, the condition in which two or more predictors are highly correlated with one another, the coefficients might be inflated and have signs that don't make sense (for example, there might be a negative coefficient when the relationship is positive.)

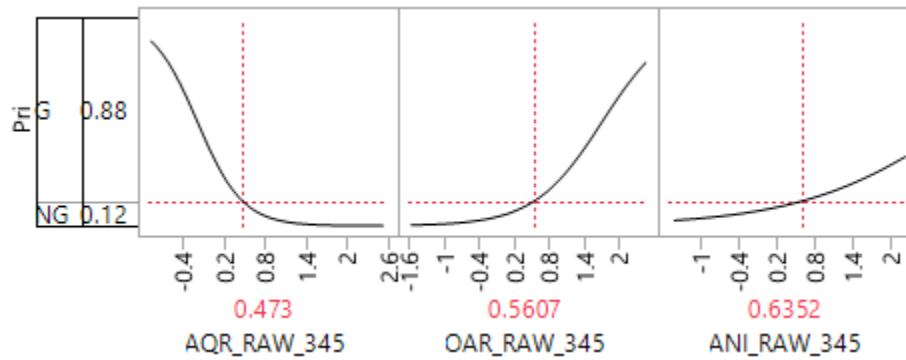


Figure 9. Logistics model prediction profiler for ASTB345

b. Bootstrap (Random) Forest Model

This model uses a bootstrap (random) forest to predict the outcome of primary training from the full set of potential predictors (main effects). Figure 10 shows the overall model summary statistics. For the validation set, R^2 is 0.0149 and the misclassification rate is 0.0876.

Measure	Training	Validation	Definition
Entropy RSquare	0.1214	0.0149	$1 - \text{Loglike}(\text{model}) / \text{Loglike}(0)$
Generalized RSquare	0.1552	0.0197	$(1 - (L(0)/L(\text{model}))^{(2/n)}) / (1 - L(0)^{(2/n)})$
Mean -Log p	0.2603	0.2925	$\sum -\text{Log}(p[j]) / n$
RMSE	0.2697	0.2814	$\sqrt{\sum (y[j] - \rho[j])^2 / n}$
Mean Abs Dev	0.1506	0.1565	$\sum y[j] - \rho[j] / n$
Misclassification Rate	0.0873	0.0876	$\sum (\rho[j] \neq p\text{Max}) / n$
N	6127	2626	n

Figure 10. Bootstrap forest overall model summary statistics for ASTB345

Figure 11 shows the column contribution of the different predictors with the z-scores for Flight Officer Flight Aptitude Rating (FOFAR_RAW_345), Spatial Awareness Rating (SAT_RAW_345), and Academic Qualifications Rating (AQR_RAW_345) being the top three contributors.

Term	Number of Splits	G ²	Portion
FOFAR_RAW_345	184	34.7870297	0.1432
SAT_RAW_345	183	28.0446458	0.1154
AQR_RAW_345	193	25.0525743	0.1031
MST_RAW_345	189	20.7710918	0.0855
OAR_RAW_345	197	20.2043049	0.0831
PFAR_RAW_345	160	18.7433502	0.0771
MCT_RAW_345	179	17.7393012	0.0730

Term	Number of Splits	G ²	Portion
RCT_RAW_345	177	17.3101868	0.0712
ANI_RAW_345	158	16.7881276	0.0691
FormalFlightInstrHours	138	15.7854347	0.0650
Race	126	14.6366291	0.0602
RecruitingBranch_A	101	5.75253092	0.0237
ExamineeStatus_A	97	4.62666663	0.0190
Gender	68	2.75067291	0.0113

Figure 11. Bootstrap forest model column contribution for ASTB345

Figure 12 shows the confusion matrix for the training and validation sets.

Training			Validation		
Actual	Predicted Count		Actual	Predicted Count	
Pri	G	NG	Pri	G	NG
G	5592	0	G	2396	0
NG	535	0	NG	230	0

Figure 12. Bootstrap forest model confusion matrix for ASTB345

As indicated earlier and as a result of the data imbalance, the model does not predict any Non-Graduates. Similar to the approach used in the logistics regression model, we specify a probability cut off lower than 0.5 to classify more observations as Non-Graduates. For example, the confusion matrix for a cut off of 0.1 which classifies the predicted probability of Non-Graduate greater than 0.10 as Non-Graduate is shown in Figure 13. The classifier of this model now predicts Non-Graduates, but with a higher overall misclassification rate.

Training			Validation			Specified Profit Matrix	
Actual	Decision Count		Actual	Decision Count		Decision	
Pri	G	NG	Pri	G	NG	Actual	NG
G	4402	1190	G	1793	603	G	-0.111
NG	133	402	NG	151	79	NG	0
Actual	Decision Rate		Actual	Decision Rate			
Pri	G	NG	Pri	G	NG		
G	0.787	0.213	G	0.748	0.252		
NG	0.249	0.751	NG	0.657	0.343		
Misclassification Rate			Misclassification Rate				
0.2159			0.2871				

Figure 13. Bootstrap forest model decision Matrix for ASTB345

Figure 14 shows the prediction profiler. It indicates the predicted probability of the response variable PRI for each specified value of the predictor variables. As shown in the figure, a higher z-score for Flight Officer Flight Aptitude Rating (FOFAR_RAW), Spatial Awareness Rating (SAT_RAW) and Academic Qualifications Rating (AQR_RAW) indicate a lower probability, albeit a small one, of Non-Graduate outcome.

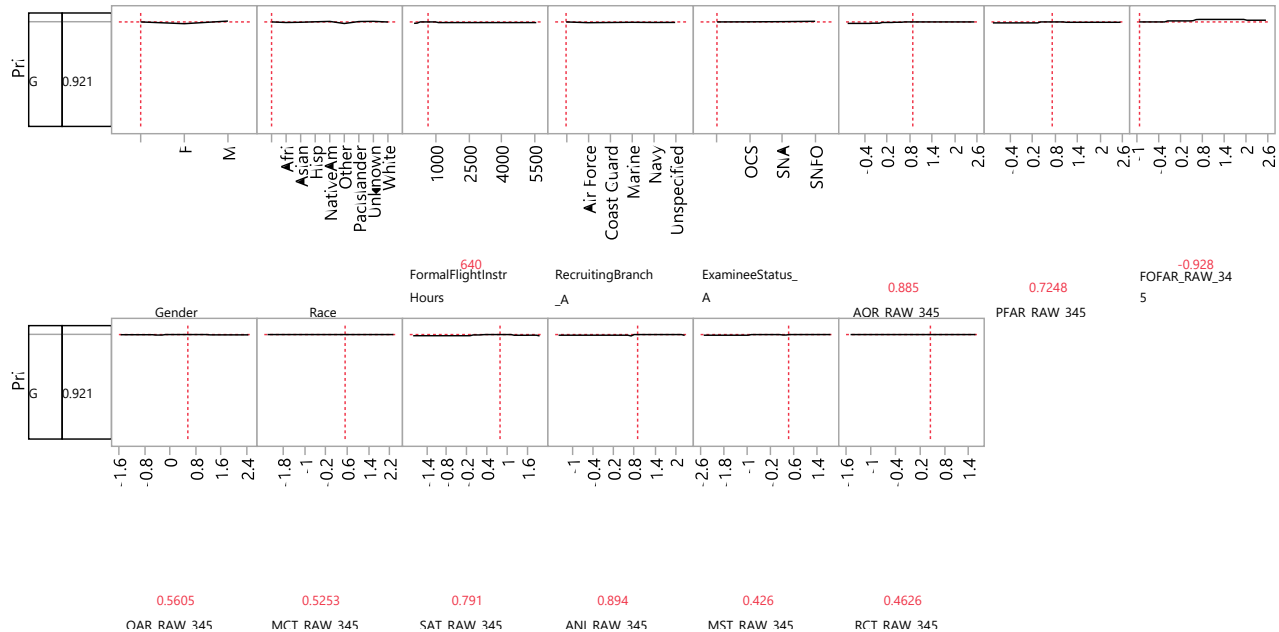


Figure 14. Bootstrap forest model prediction profiler for ASTB345

c. Penalized Regression Models

To address the issue of multicollinearity of some predictor variables, we fit several penalized generalized regression models. These models include Lasso, Elastic Net, and Ridge. Figure 15 shows the result of the Lasso Model. Note that several parameter coefficients were shrunk to zero to address the problems of multicollinearity.

Model Summary

Response	Pri
Distribution	Binomial
Estimation Method	Lasso
Validation Method	Validation Column
Probability Model Link	Logit

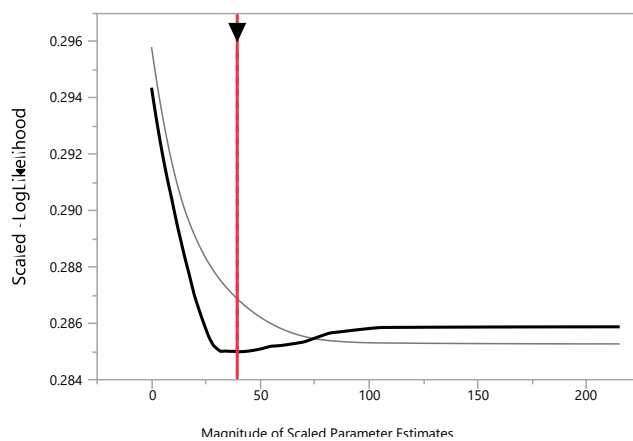
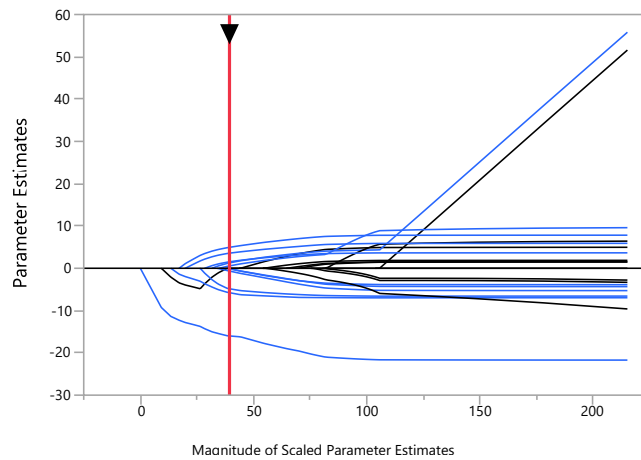
Measure	Training	Validation
---------	----------	------------

Measure	Training	Validation
Number of rows	3593	1526
Sum of Frequencies	3593	1526
-LogLikelihood	1030.7316	434.89579
Number of Parameters	12	12
BIC	2159.7041	957.75645
AICc	2085.5504	893.9978
Generalized RSquare	0.0396687	0.0417566
Lambda Penalty	0.2641608	.

Estimation Details

Number of Grid Points 150
Minimum Penalty Fraction 1e-4
Grid Scale Square Root

Solution Path



Legend

— Validation
— Training

Parameter Estimates for Original Predictors

Term	Estimate	Std Error	Wald ChiSquare	Prob > ChiSquare	Lower 95%	Upper 95%
Intercept	-1.852985	0.1283422	208.45141	<.0001*	-2.104531	-1.601439
FOFAR_RAW_345	-0.495673	0.1752393	8.0006923	0.0047*	-0.839136	-0.15221
Race[Asian-White]	0.4817107	0.315824	2.3263892	0.1272	-0.137293	1.1007142
SAT_RAW_345	-0.155737	0.127855	1.483701	0.2232	-0.406328	0.0948545
ANI_RAW_345	-0.145208	0.1283487	1.2799584	0.2579	-0.396766	0.1063512
Race[Afri-White]	0.3523081	0.317655	1.2300814	0.2674	-0.270284	0.9749006
Race[Other-White]	0.3956528	0.954941	0.1716625	0.6786	-1.475997	2.2673028
RecruitingBranch_A[Marine-Unspecified]	0.0534161	0.1574944	0.1150308	0.7345	-0.255267	0.3620996
FormalFlightInstrHours	-0.000052	0.0002677	0.0376482	0.8462	-0.000577	0.0004728
ExamineeStatus_A[SNA-SNFO]	0.0347326	0.2297834	0.0228474	0.8799	-0.415635	0.4850997
Race[PacIslander-White]	-0.032552	0.5045994	0.0041617	0.9486	-1.021549	0.9564442
Gender[F-M]	-0.002559	0.1941245	0.0001737	0.9895	-0.383036	0.3779186
AQR_RAW_345	0	0	0	1.0000	0	0
PFAR_RAW_345	0	0	0	1.0000	0	0
OAR_RAW_345	0	0	0	1.0000	0	0
MCT_RAW_345	0	0	0	1.0000	0	0
MST_RAW_345	0	0	0	1.0000	0	0
RCT_RAW_345	0	0	0	1.0000	0	0
RecruitingBranch_A[Air Force-Unspecified]	0	0	0	1.0000	0	0
RecruitingBranch_A[Coast Guard-Unspecified]	0	0	0	1.0000	0	0
RecruitingBranch_A[Navy-Unspecified]	0	0	0	1.0000	0	0
ExamineeStatus_A[OCS-SNFO]	0	0	0	1.0000	0	0
Race[Hisp-White]	0	0	0	1.0000	0	0
Race[NativeAm-White]	0	0	0	1.0000	0	0
Race[Unknown-White]	0	0	0	1.0000	0	0

Figure 15. Lasso model for ASTB345

Figure 16 shows the prediction profiler with confidence intervals.

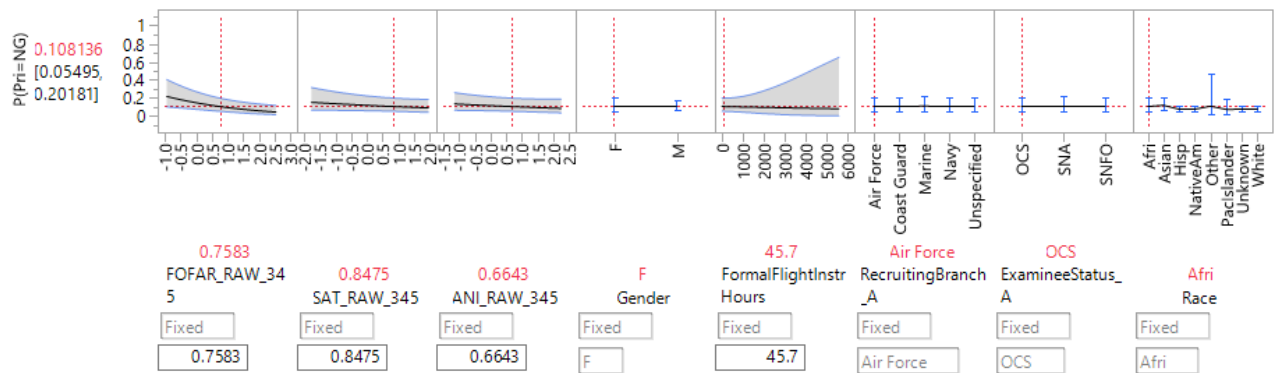


Figure 16. Lasso model prediction profiler for ASTB345

Figure 17 shows the result of the Elastic Net Model. Similar to the Lasso Model, the Elastic Net model shrank the coefficients of some the parameters to zero to address the problems of multicollinearity.

Model Summary

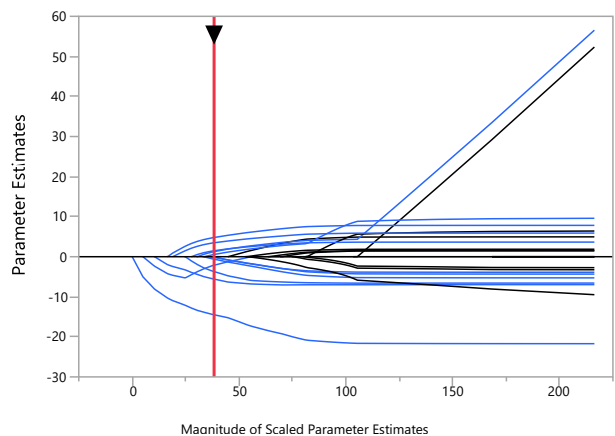
Response	Pri
Distribution	Binomial
Estimation Method	Elastic Net
Validation Method	Validation Column
Probability Model Link	Logit

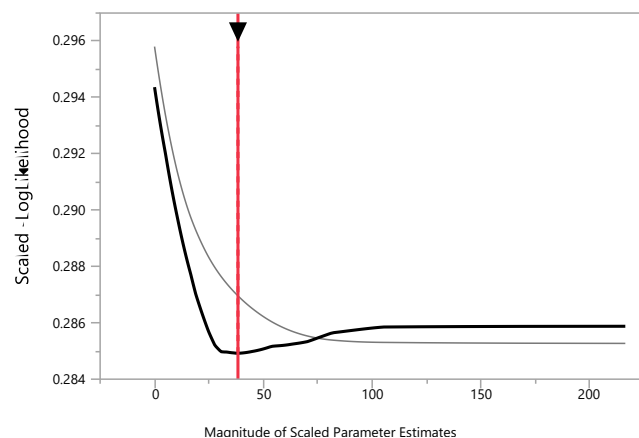
Measure	Training	Validation
Number of rows	3593	1526
Sum of Frequencies	3593	1526
-LogLikelihood	1031.0422	434.78611
Number of Parameters	13	13
BIC	2168.512	964.86749
AICc	2088.1861	895.81296
Generalized RSquare	0.0392884	0.0420736
Lambda Penalty	0.2668291	.

Estimation Details

Elastic Net Alpha	0.99
Number of Grid Points	150
Minimum Penalty Fraction	1e-4
Grid Scale	Square Root

Solution Path





Legend

— Validation

— Training

Parameter Estimates for Original Predictors

Term	Estimate	Std Error	Wald ChiSquare	Prob > ChiSquare	Lower 95%	Upper 95%
Intercept	-1.858633	0.128437	209.41464	<.0001*	-2.110365	-1.606901
FOFAR_RAW_345	-0.449022	0.2136298	4.4178487	0.0356*	-0.867728	-0.030315
Race[Asian-White]	0.4662487	0.3192925	2.132347	0.1442	-0.159553	1.0920506
Race[Afri-White]	0.3430654	0.3182929	1.1617156	0.2811	-0.280777	0.9669081
SAT_RAW_345	-0.150579	0.145617	1.0693137	0.3011	-0.435983	0.134825
ANI_RAW_345	-0.111317	0.2543582	0.1915268	0.6616	-0.60985	0.3872161
Race[Other-White]	0.3851169	0.9601999	0.1608651	0.6884	-1.49684	2.267074
RecruitingBranch_A[Marine-Unspecified]	0.0534714	0.1578054	0.1148151	0.7347	-0.255822	0.3627643
FormalFlightInstrHours	-0.000048	0.0002663	0.0325214	0.8569	-0.00057	0.0004739
PFAR_RAW_345	-0.07028	0.4011843	0.0306884	0.8609	-0.856587	0.7160269
ExamineeStatus_A[SNA-SNFO]	0.0341913	0.2299371	0.0221113	0.8818	-0.416477	0.4848598
Race[PacIslander-White]	-0.030503	0.5049068	0.0036498	0.9518	-1.020102	0.9590961
Gender[F-M]	-0.004096	0.1944628	0.0004436	0.9832	-0.385236	0.3770444
AQR_RAW_345	0	0	0	1.0000	0	0
ExamineeStatus_A[OCS-SNFO]	0	0	0	1.0000	0	0
MCT_RAW_345	0	0	0	1.0000	0	0
MST_RAW_345	0	0	0	1.0000	0	0
OAR_RAW_345	0	0	0	1.0000	0	0
Race[Hisp-White]	0	0	0	1.0000	0	0
Race[NativeAm-White]	0	0	0	1.0000	0	0
Race[Unknown-White]	0	0	0	1.0000	0	0
RCT_RAW_345	0	0	0	1.0000	0	0
RecruitingBranch_A[Air Force-Unspecified]	0	0	0	1.0000	0	0
RecruitingBranch_A[Coast Guard-Unspecified]	0	0	0	1.0000	0	0
RecruitingBranch_A[Navy-Unspecified]	0	0	0	1.0000	0	0

Figure 17. Elastic net model for ASTB345

Figure 18 shows the prediction profiler with confidence intervals for the Elastic Net Model.

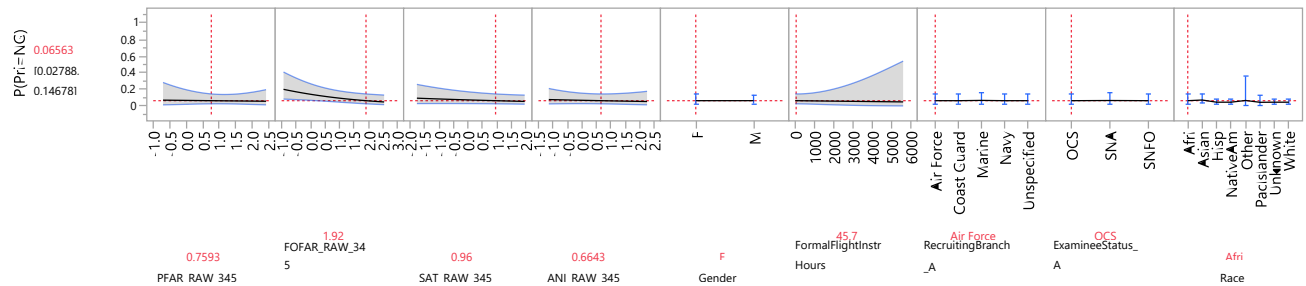


Figure 18. Elastic net model prediction profiler for ASTB345

Figure 19 shows the result of the Ridge Model. As shown in the Solutions Path graph and unlike the Lasso and Elastic Net models, the coefficient parameters do not shrink to zero and the final model includes all the coefficients.

Model Summary

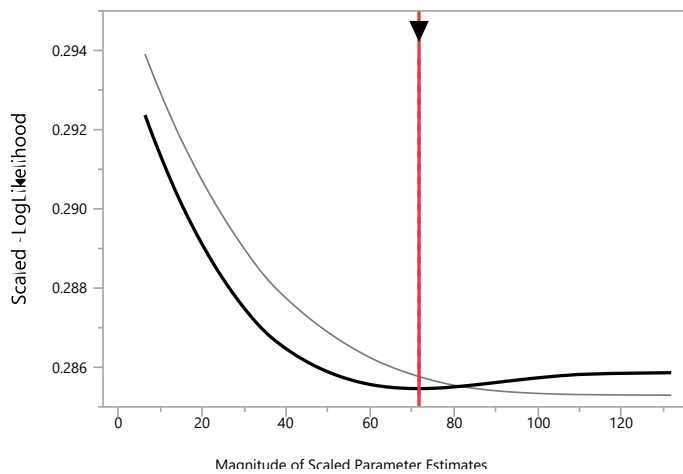
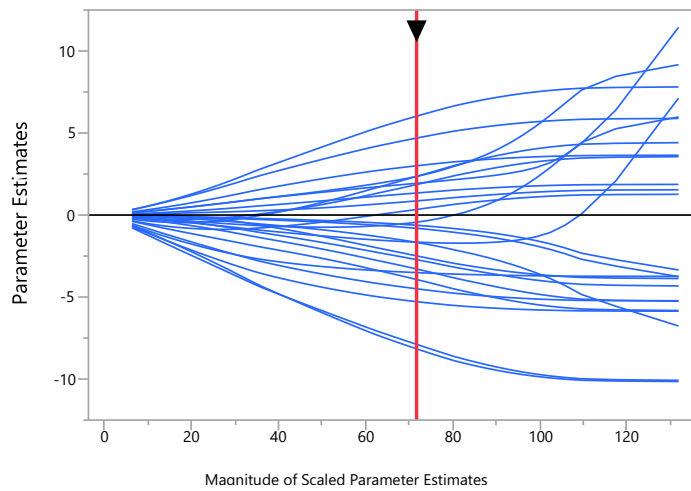
Response	Pri
Distribution	Binomial
Estimation Method	Ridge
Validation Method	Validation Column
Probability Model Link	Logit

Measure	Training	Validation
Number of rows	3593	1526
Sum of Frequencies	3593	1526
-LogLikelihood	1026.7551	435.61374
Number of Parameters	25	25
BIC	2258.1788	1054.4876
AICc	2103.8747	922.09415
Generalized RSquare	0.0445321	0.0396803
Lambda Penalty	0.0291936	.

Estimation Details

Number of Grid Points	150
Minimum Penalty Fraction	1e-4
Grid Scale	Square Root

Solution Path



Legend

— Validation
— Training

Parameter Estimates for Original Predictors

Term	Estimate	Std Error	Wald ChiSquare	Prob > ChiSquare	Lower 95%	Upper 95%
Race[Asian-White]	0.5862937	0.3023869	3.7592783	0.0525	-0.006374	1.1789611
Race[Unknown-White]	-1.180045	0.7594321	2.4144543	0.1202	-2.668504	0.3084149
Race[Afri-White]	0.4626728	0.2980507	2.409725	0.1206	-0.121496	1.0468415
Race[Other-White]	0.776877	0.7625107	1.0380365	0.3083	-0.717617	2.2713705
Race[PacIslander-White]	-0.557376	0.7680206	0.5266856	0.4680	-2.062669	0.9479162
Gender[F-M]	-0.143986	0.2081716	0.478405	0.4891	-0.551994	0.2640231
FormalFlightInstrHours	-0.000258	0.0003746	0.4737685	0.4913	-0.000992	0.0004763
RecruitingBranch_A[Coast Guard- Unspecified]	-0.810371	1.2270079	0.4361871	0.5090	-3.215262	1.5945205
RecruitingBranch_A[Air Force- Unspecified]	-0.624282	1.2326062	0.2565154	0.6125	-3.040146	1.7915814
Race[NativeAm-White]	0.3362463	0.8010344	0.1762027	0.6747	-1.233752	1.9062448

Term	Estimate	Std Error	Wald ChiSquare	Prob > ChiSquare	Lower 95%	Upper 95%
ExamineeStatus_A[SNA-SNFO]	0.1452356	0.5727346	0.0643042	0.7998	-0.977304	1.2677747
Race[Hisp-White]	0.0706385	0.2832797	0.0621801	0.8031	-0.48458	0.6258565
RecruitingBranch_A[Marine-Unspecified]	0.0814749	0.6778086	0.0144489	0.9043	-1.247006	1.4099555
RecruitingBranch_A[Navy-Unspecified]	-0.06939	0.6688172	0.0107643	0.9174	-1.380248	1.2414672
ExamineeStatus_A[OCS-SNFO]	-0.025717	0.5395349	0.002272	0.9620	-1.083186	1.0317519
Intercept	-1.752643	12041.546	2.1185e-8	0.9999	-23602.75	23599.244
SAT_RAW_345	-0.220038	27092.516	6.596e-11	1.0000	-53100.58	53100.135
RCT_RAW_345	0.056829	13280.708	1.831e-11	1.0000	-26029.65	26029.767
FOFAR_RAW_345	-0.244186	61597.251	1.572e-11	1.0000	-120728.6	120728.15
ANI_RAW_345	-0.134668	52353.961	6.617e-12	1.0000	-102612	102611.74
PFAR_RAW_345	-0.175625	77470.029	5.139e-12	1.0000	-151838.6	151838.29
MST_RAW_345	-0.095792	44695.035	4.593e-12	1.0000	-87600.76	87600.564
AQR_RAW_345	-0.111204	92919.545	1.432e-12	1.0000	-182119.1	182118.85
MCT_RAW_345	0.0635753	63784.603	9.934e-13	1.0000	-125015.5	125015.59
OAR_RAW_345	0.0093435	107393.22	7.569e-15	1.0000	-210486.8	210486.84

Figure 19. Ridge model for ASTB345

Figure 20 shows the prediction profiler with confidence intervals for the Ridge Model.

Prediction Profiler

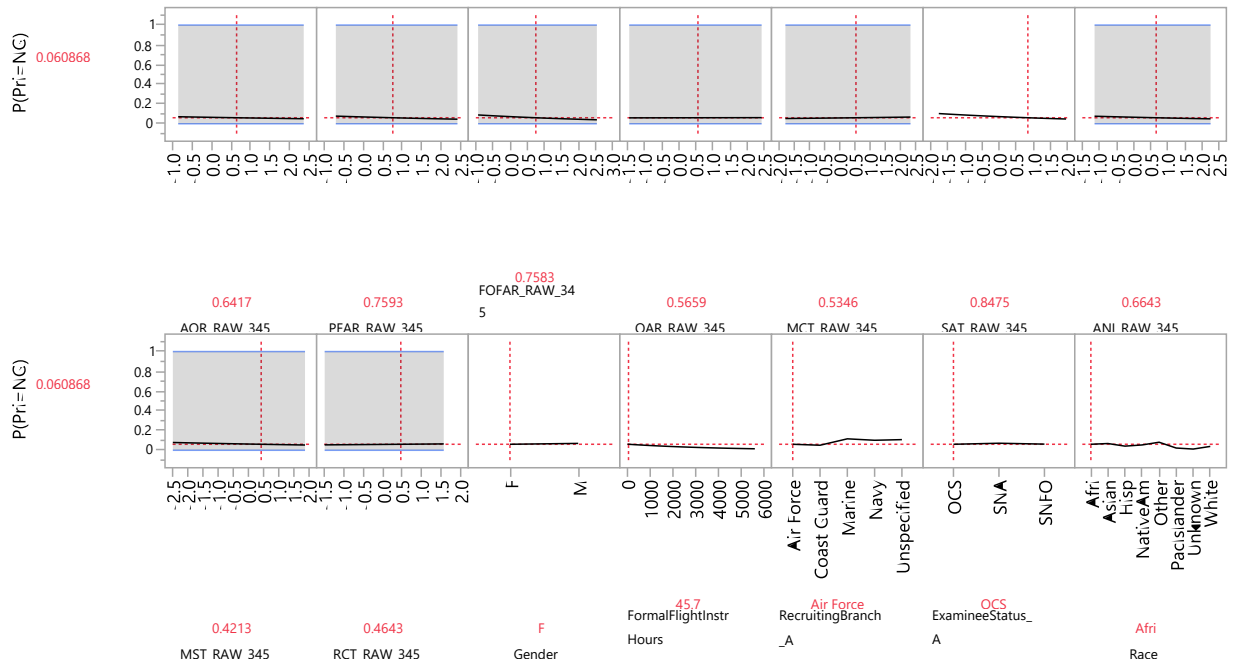


Figure 20. Ridge model prediction profiler for ASTB345

d. Model Comparison

We compare the six developed models to select a best model. Figure 21 and Figure 22 compare these models for both the training and validation data sets. As

shown in these figures, the Bootstrap Forrest Model has the best R^2 for the training data set, while the Elastic Net has the worst. This situation is reversed for the validation set. This is not surprising since decision trees tend to overfit the data and hence have a good performance on the training data set. On the other hand, Elastic Net as a penalized model would generalize well on validation and test data sets.

Creator		Entropy RSquare	Generalize d RSquare	Mean - Log p	RMSE	Mean Abs Dev	Misclassification Rate	N
Bootstrap Forest		0.1214	0.1552	0.2603	0.2697	0.1506	0.0873	6127
Fit Nominal Logistic		0.0403	0.0528	0.2839	0.2786	0.1552	0.0871	3593
Fit Generalized Ridge		0.0340	0.0445	0.2858	0.2791	0.1562	0.0871	3593
Fit Nominal Logistic		0.0321	0.0421	0.2869	0.2794	0.1562	0.0873	6125
Fit Generalized Lasso		0.0302	0.0397	0.2869	0.2794	0.1566	0.0871	3593
Fit Generalized Elastic Net		0.0299	0.0393	0.287	0.2794	0.1567	0.0871	3593

Figure 21. Model comparison on training data set for ASTB345

Creator		Entropy RSquare	Generalize d RSquare	Mean - Log p	RMSE	Mean Abs Dev	Misclassification Rate	N
Fit Generalized Elastic Net		0.0321	0.0421	0.2849	0.2781	0.1544	0.0865	1526
Fit Generalized Lasso		0.0318	0.0417	0.285	0.2781	0.1544	0.0865	1526
Fit Generalized Ridge		0.0303	0.0397	0.2855	0.2784	0.1543	0.0865	1526
Fit Nominal Logistic		0.0264	0.0347	0.2866	0.2786	0.1536	0.0865	1526
Fit Nominal Logistic		0.0183	0.0242	0.2915	0.2814	0.1564	0.0876	2625
Bootstrap Forest		0.0149	0.0197	0.2925	0.2814	0.1565	0.0876	2626

Figure 22. Model comparison on validation data set for ASTB345

e. Conclusions

In this section, we developed a multiple logistics model, a bootstrap (random) forest, and several penalized regression models (Lasso, Elastic Net, and Ridge) to predict the outcome of primary training (PRI) (Graduate or Non-Graduate/Attrite) from ASTB345 data. All models developed in this section do not seem to provide a good fit as evident by the small R^2 values of the validation set (<4%), which indicate that only a small percentage of the dependent variable variation is explained by the model.

- The most important predictor variables seem to be the Z-score for the Flight Officer Flight Aptitude Rating (FOFAR_RAW_345).

- Demographic variables don't seem to be significant in predicting the outcome variable.
- Due to collinearity among predictor variables, tree and penalized models would be more accurate models. In fact, model comparison on validation data indicated the three generalized linear models performed the best.

2. Predicting Outcome of Primary Training (PRI) from IFS Scores

In this section, we develop multiple logistics models, a bootstrap (random) forest model, and several penalized general regression models (Lasso, Elastic Net, and Ridge) to predict the outcome of primary training (PRI) (Graduate or Non-Graduate/Attrite) from IFS data. IFS Predictors include number of hours of flight time received during IFS, number of landings completed in IFS, exam score on Stages 1, 2, and 3 Exams, End of Course exam score, FAA exam score, and number of flight failures in IFS. We note that there exists some correlation between the IFS exam scores, as shown in Figure 23.

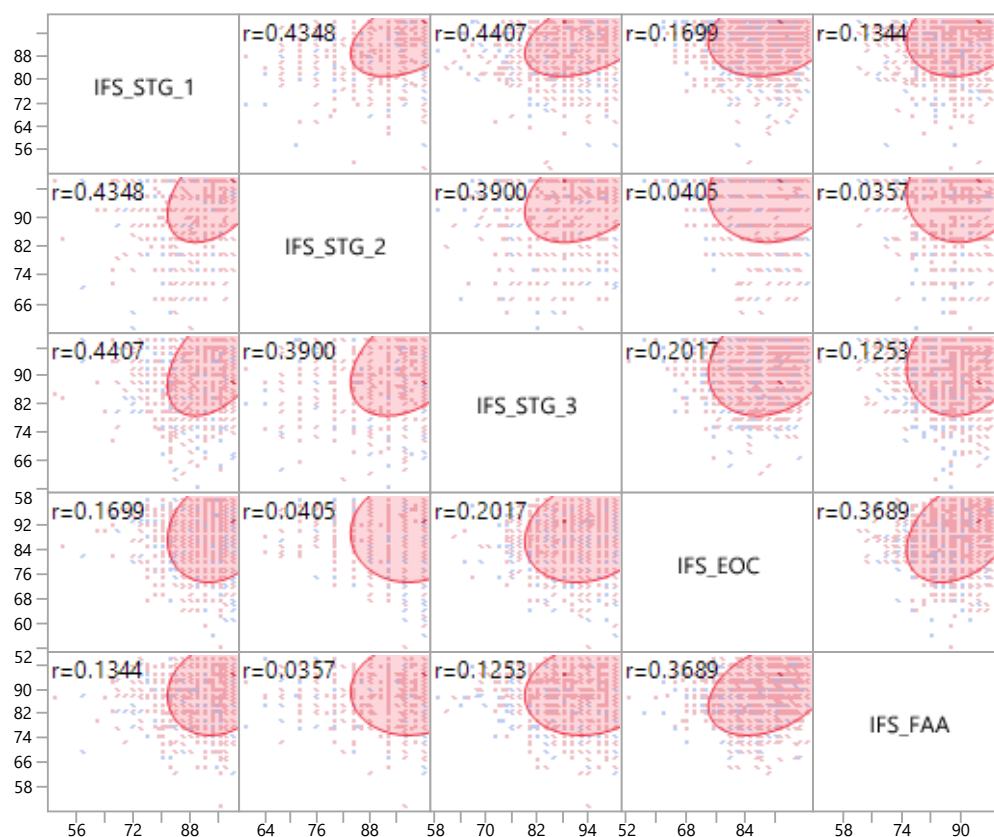


Figure 23. Correlation of IFS scores

a. Multiple Logistics Model with Backward Elimination

This multiple logistics model predicts the outcome of primary training from the full set of potential predictors (main effects). The model is then reduced using backward elimination to remove non-significant predictors. Figure 24 shows the effect summary of the model after removing the insignificant predictors.

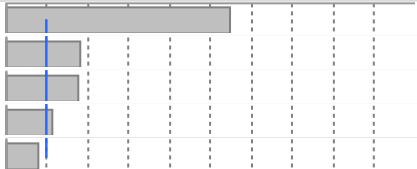
Source	LogWorth		PValue
IFS_FLT_FAIL	10.946		0.00000
IFS_FAA	3.661		0.00022
IFS_STG_1	3.578		0.00026
IFS_TOTAL_LANDINGS	2.275		0.00531
IFS_EOC	1.652		0.02229

Figure 24. Logistics model effect summary for IFS

As shown in Figure 24, the top three significant effects for predicting the outcome of primary training are the number of flight failures in IFS, the FAA exam score, and exam score on Stage 1. The Whole Model Test Report, shown in Figure 25, has a small p-value (<0.0001), which indicates that the overall model is significant, and an R^2 of 0.0534 indicating that 5.34 % of the variation in response is explained by the model.

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	53.20791	5	106.4158	$<.0001^*$
Full	942.96899			
Reduced	996.17690			
	RSquare (U)		0.0534	
	AICc		1897.97	
	BIC		1934	
	Observations (or Sum Wgts)		3012	

Figure 25. Logistics model whole model test

Figure 26 shows the confusion matrix for both the training and validation sets.

Training			Validation		
Actual	Predicted Count		Actual	Predicted Count	
Pri	NG	G	Pri	NG	G
NG	1	308	NG	1	134
G	5	2698	G	1	1152

Figure 26. Logistics model confusion matrix for IFS

As a result of the data imbalance, the model does not predict any Non-Graduates. To remedy this situation, we specify a cut off lower than 0.5 needs to be specified for classifying a predicted probability as Non-Graduates. For example, the confusion matrix for a cut off of 0.1 is shown in Figure 27. The classifier of this model now predicts Non-Graduates but with a higher overall misclassification rate.

Training			Validation			Specified Profit Matrix		
Actual			Actual			Decision		
Pri	NG	G	Pri	NG	G	Actual	NG	G
NG	181	128	NG	68	67	NG	0	-1
G	931	1772	G	404	749	G	-0.111	0
Actual			Actual					
Pri	NG	G	Pri	NG	G			
NG	0.586	0.414	NG	0.504	0.496			
G	0.344	0.656	G	0.350	0.650			
Misclassification Rate			Misclassification Rate					
0.3516			0.3657					

Figure 27. Logistics model decision matrix for IFS

Figure 28 shows the prediction profiler. It indicates the predicted probability of the response variable PRI for each specified value of the predictor variables. As shown in the figure, a higher IFS Stage 1 Score, IFS End of Course Score, IFS FAA Score, and IFS total landings indicate a lower probability of Non-Graduate outcome, while a higher number of IFS flight failures is associated with a higher probability of Non-Graduate.

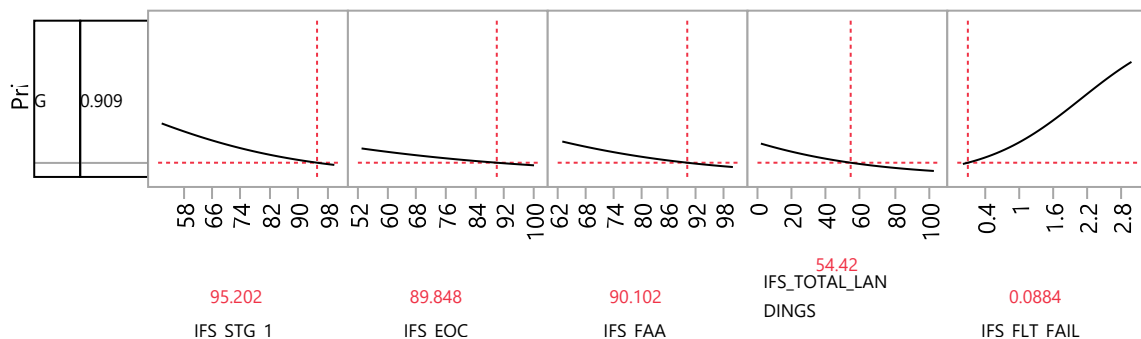


Figure 28. Logistics model prediction profiler for IFS

b. Bootstrap (Random) Forest Model

This model uses a bootstrap (random) forest to predict the outcome of primary training from the full set of potential predictors (main effects). Figure 29 shows the overall model summary statistics. For the validation set, R^2 is 0.0135 and the misclassification rate is 0.0876.

Measure	Training	Validation	Definition
Entropy RSquare	0.2015	0.0135	$1 - \text{Loglike}(\text{model}) / \text{Loglike}(0)$
Generalized RSquare	0.2517	0.0178	$(1 - (L(0)/L(\text{model}))^{2/n}) / (1 - L(0)^{2/n})$
Mean -Log p	0.2366	0.2929	$\sum -\log(p[j]) / n$
RMSE	0.2579	0.2815	$\sqrt{\sum (y[j] - p[j])^2 / n}$
Mean Abs Dev	0.1419	0.1528	$\sum y[j] - p[j] / n$
Misclassification Rate	0.0873	0.0876	$\sum (p[j] \neq pMax) / n$
N	6127	2626	n

Figure 29. Bootstrap forest overall model summary for IFS

Figure 30 shows the column contribution of the different predictors with hours of flight time received during IFS (IFS_TOTAL_FLIGHT_TIME), FAA Exam Score (IFS_FAA), and number of landings completed during IFS (IFS_TOTAL_LANDINGS) being the top three contributors.

Term	Number of Splits	G ²	Portion
IFS_TOTAL_FLIGHT_TIME	2246	81.3450911	0.1683
IFS_FAA	2137	81.1402341	0.1679
IFS_TOTAL_LANDINGS	2278	80.4154733	0.1664
IFS_STG_3	2087	69.7620513	0.1443
IFS_STG_1	1955	58.4733708	0.1210
IFS_EOC	1771	55.6287247	0.1151
IFS_STG_2	1444	30.2487696	0.0626
IFS_FLT_FAIL	296	26.3946661	0.0546

Figure 30. Bootstrap forest model column contribution for IFS

Figure 31 shows the confusion matrix for both the training and validation sets.

Training			Validation		
Actual	Predicted Count		Actual	Predicted Count	
Pri	G	NG	Pri	G	NG

G	5592	0	G	2396	0
NG	535	0	NG	230	0

Figure 31. Bootstrap forest model confusion matrix for IFS

As indicated earlier and as a result of the data imbalance, although the misclassification rate is low, the model does not predict any Non-Graduates. Similar to the approach used in the logistics regression model, we need to specify a probability cut off lower than 0.5 to classify more observations as Non-Graduates.

Figure 32 shows the prediction profiler. It indicates the predicted probability of the response variable PRI for each specified value of the predictor variables. As shown in Figure 32, a higher score in IFS exams is associated with a lower probability of Non-Graduate outcome, while a higher number of flight fails (IFS_FLT_FAIL) is associated with a higher probability of Non-Graduate outcome.

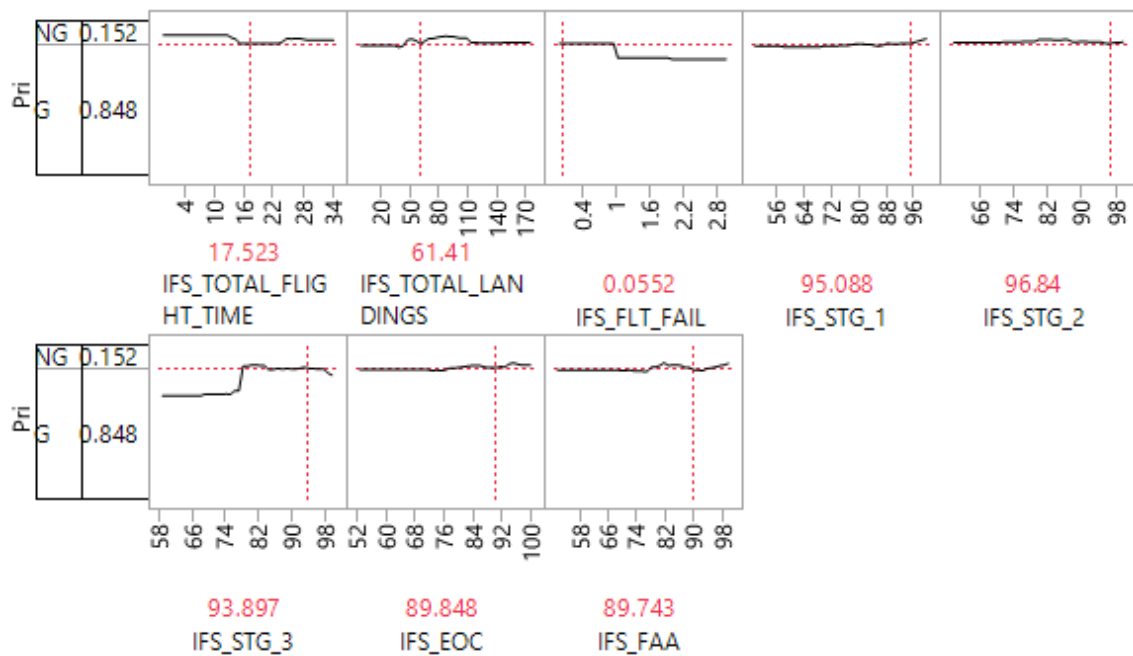


Figure 32. Bootstrap forest model prediction profiler for IFS

c. Penalized Regression Model

To address the issue of multicollinearity of some predictor variables, we fit a penalized generalized regression model. We choose to fit an Elastic Net model since this model performed slightly better than other penalized models in previous analyses.

Figure 33 shows the result of the model. Note that several parameter coefficients were shrunk to zero to address the problems of multicollinearity.

Model Summary

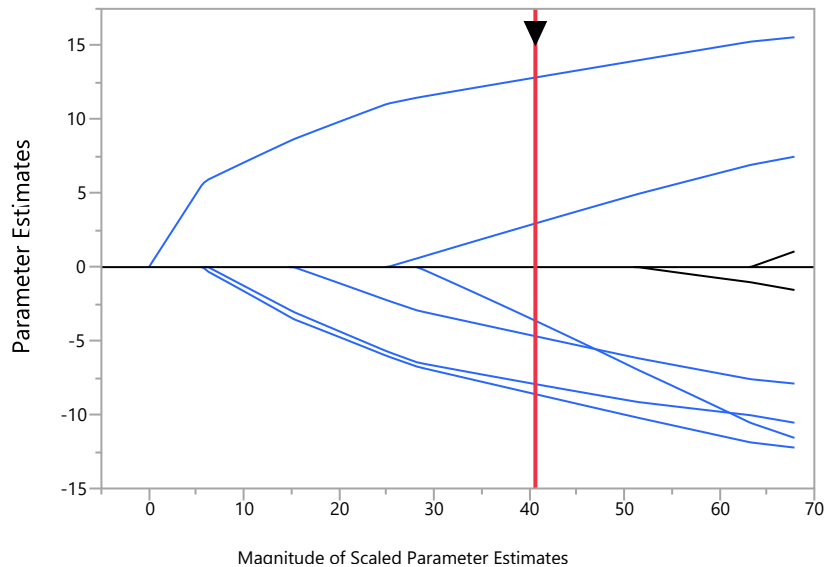
Response	Pri
Distribution	Binomial
Estimation Method	Elastic Net
Validation Method	Validation Column
Probability Model Link	Logit

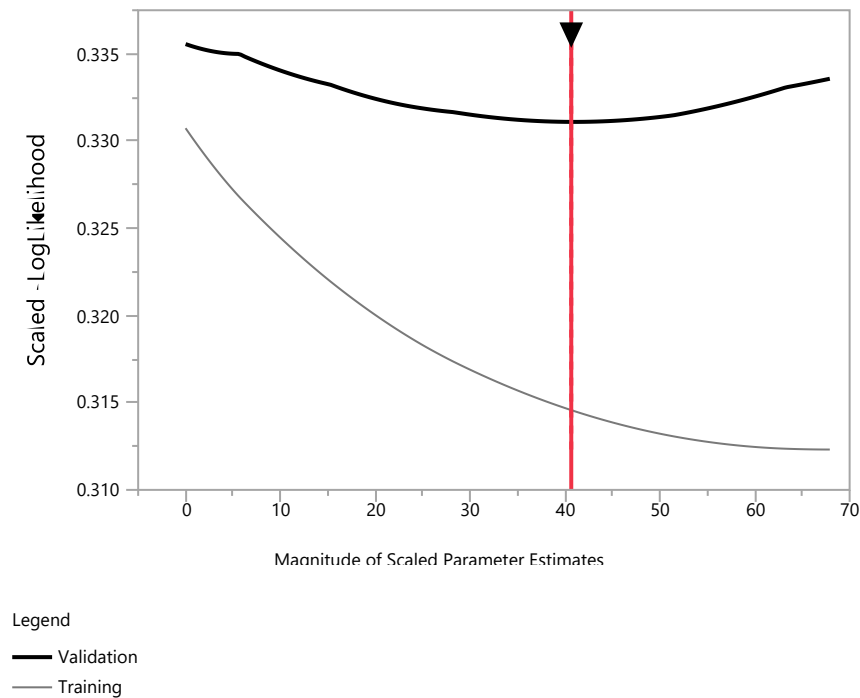
Measure	Training	Validation
Number of rows	3012	1288
Sum of Frequencies	3012	1288
-LogLikelihood	947.49399	426.45516
Number of Parameters	7	7
BIC	1951.0605	903.03625
AICc	1909.0253	866.99783
Generalized RSquare	0.0657336	0.0181681
Lambda Penalty	0.5124345	.

Estimation Details

Elastic Net Alpha	0.99
Number of Grid Points	150
Minimum Penalty Fraction	1e-4
Grid Scale	Square Root

Solution Path





Parameter Estimates for Original Predictors

Term	Estimate	Std Error	Wald ChiSquare	Prob > ChiSquare	Lower 95%	Upper 95%
IFS_FLT_FAIL	0.7878577	0.1882748	17.511004	<.0001*	0.4188458	1.1568696
IFS_STG_1	-0.026937	0.0107396	6.290901	0.0121*	-0.047986	-0.005887
IFS_FAA	-0.027076	0.0109557	6.1080574	0.0135*	-0.048549	-0.005604
Intercept	3.4876873	1.6692871	4.3652891	0.0367*	0.2159447	6.7594299
IFS_EOC	-0.012879	0.0098199	1.7199887	0.1897	-0.032125	0.006368
IFS_TOTAL_LANDINGS	-0.006942	0.0066424	1.092334	0.2960	-0.019961	0.0060766
IFS_TOTAL_FLIGHT_TIME	0.0539102	0.0589162	0.8372849	0.3602	-0.061563	0.1693838
IFS_STG_2	0	0	0	1.0000	0	0
IFS_STG_3	0	0	0	1.0000	0	0

Figure 33. Elastic net model for IFS

As shown in Figure 34, the most significant predictors are the number of flight failures in IFS (IFS_FLT_FAIL), Exam Score on Stage 1 Exam (IFS_STG_1), and FAA Exam Score (IFS_FAA). Figure 34 shows the prediction profiler. Higher values of exam scores (IFS_STG_1, IFS_EOC, IFS_FAA) and number of landings completed in IFS (IFS_TOTAL_LANDINGS) are associated with lower probability of Non-Graduate outcome, while higher numbers of flight failures are associated with higher probability of Non-Graduate outcome.

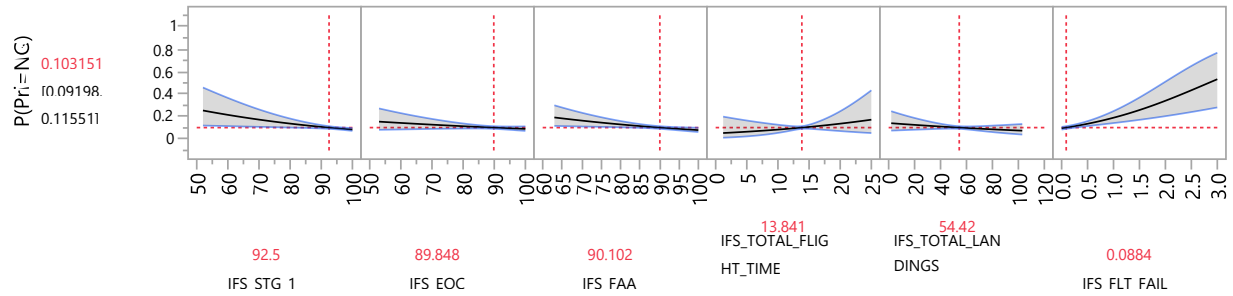


Figure 34. Elastic net model prediction profiler for IFS

d. Model Comparison

We compare the three developed models to select a best model. Figure 35 and Figure 36 compare these models for both the training and validation data sets. As shown in these figures, the Bootstrap Forrest Model has the best R^2 for the training data set, while the Elastic Net has the worst. This situation is reversed for the validation set. As discussed earlier, this is not surprising since decision trees tend to overfit the data and hence have a good performance on the training data set. On the other hand, Elastic Net as a penalized model would generalize well on validation and test data sets.

Creator		Entropy RSquare	Generalized RSquare	Mean - Log p	RMSE	Mean Abs Dev	Misclassification Rate	N
Bootstrap Forest		0.2009	0.2511	0.2368	0.2582	0.1417	0.0873	6127
Fit Nominal Logistic		0.0534	0.0717	0.3131	0.2973	0.1765	0.1039	3012
Fit Generalized Elastic Net		0.0489	0.0657	0.3146	0.2978	0.1788	0.1029	3012

Figure 35. Model comparison for training set for IFS

Creator		Entropy RSquare	Generalized RSquare	Mean - Log p	RMSE	Mean Abs Dev	Misclassification Rate	N
Fit Generalized Elastic Net		0.0132	0.0181	0.3311	0.3054	0.1830	0.1056	1288
Bootstrap Forest		0.0121	0.0159	0.2933	0.2817	0.1526	0.0876	2626
Fit Nominal Logistic		0.0060	0.0082	0.3335	0.3067	0.1822	0.1048	1288

Figure 36. Model comparison for validation set for IFS

e. Conclusions

In this section, we developed a multiple logistics model, a bootstrap (random) forest model, and an Elastic Net penalized regression models to predict the outcome of

primary training (PRI) (Graduate or Non-Graduate/Attrite) from IFS data. All models developed in this section do not seem to be a good fit as evident by the small R^2 values of the validation set (<2%), which indicate that only a small percentage of the dependent variable variation is explained by the model. Based on the analysis of the developed models, we make the following conclusions:

- The top three significant effects in the logistics model for predicting the outcome are the number of flight failures in IFS, the FAA exam score, and exam score on Stage 1.
- For the bootstrap forest model, the top three variable contributors are the number of hours of flight time received during IFS (IFS_TOTAL_FLIGHT_TIME), FAA Exam Score (IFS_FAA), and number of landings completed during IFS (IFS_TOTAL_LANDINGS).
- The most significant predictors in the Elastic Net model are the number of flight failures in IFS (IFS_FLT_FAIL), Exam Score on Stage 1 Exam (IFS_STG_1), and FAA Exam Score (IFS_FAA).
- The Elastic Net model performed the best on validation data as indicated by the R^2 score.

3. Predicting Outcome of Primary Training (PRI) from API Scores

In this section, we develop a multiple logistics model and a decision (partition) tree models to predict the outcome of primary training (PRI) (Graduate or Non-Graduate/Attrite) from API data. AI Predictors include the overall performance of API as Navy Standard Score and number of academic failures in API.

a. Multiple Logistics Model

This multiple logistics model predicts the outcome of primary training from the set of potential predictors (main effects). Figure 37 shows the effect summary of the model.

Source	LogWorth		PValue
API_NSS	26.244		0.00000
API_Test_FAILS	0.844		0.14327

Figure 37. Logistics model effect summary for API

As shown in Figure 37, the most significant effect for predicting the outcome of primary training is the overall performance of API as Navy Standard Score (API_NSS). The Whole Model Test Report, shown in Figure 38, has a small p-value (<0.0001), which indicates that the overall model is significant, and an R^2 of 0.0707 indicating that about 7 % of the variation in response is explained by the model.

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	128.3332	4	256.6663	$<.0001^*$
Full	1686.9403			
Reduced	1815.2734			
RSquare (U)			0.0707	
AICc			3383.89	
BIC			3417.48	
Observations (or Sum Wgts)			6126	

Figure 38. Logistics model whole model test for API

Figure 39 shows the confusion matrix for both the training and validation sets.

Training			Validation		
Actual	Predicted Count		Actual	Predicted Count	
Pri	NG	G	Pri	NG	G
NG	2	533	NG	1	229
G	2	5589	G	0	2396

Figure 39. Logistics model confusion matrix for API

As discussed earlier and as a result of the data imbalance, the model does not predict any Non-Graduates. To remedy this situation, we need to specify a cut off lower than 0.5 for classifying a predicted probability as Non-Graduates.

Figure 40 shows the prediction profiler. It indicates the predicted probability of the response variable PRI for each value of the API predictor variables. As shown in the figure, a higher value of the overall performance of API as Navy Standard score (API_NSS) indicate a lower probability of Non-Graduate outcome, while a higher number of academic failures in API (API_Test_FAILS) is associated with a higher probability of Non-Graduate.

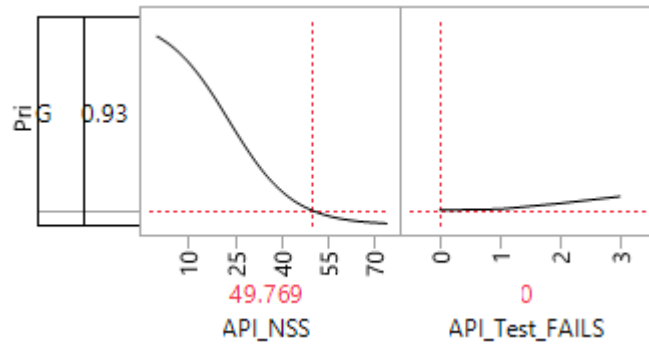


Figure 40. Logistics model prediction profiler for API

b. Decision (Partition) Tree

In this section, we fit a decision tree model to the API data. Figure 41. Decision tree shows the resulting decision tree pruned using the validation data to avoid overfitting. Most of the splits occur on the API Navy Standard Score (API_NSS), which is the most significant predictor, as show in Figure 42. Predictor contribution. The fit details and confusion matrices are shown in Figure 43, and the leaf report, shown in Figure 44, displays the rules generated by the decision tree. As indicated by the first rule of the figure, a Navy Standard Score of less than 33 is associated with a high probability of Non-Graduate Outcome.

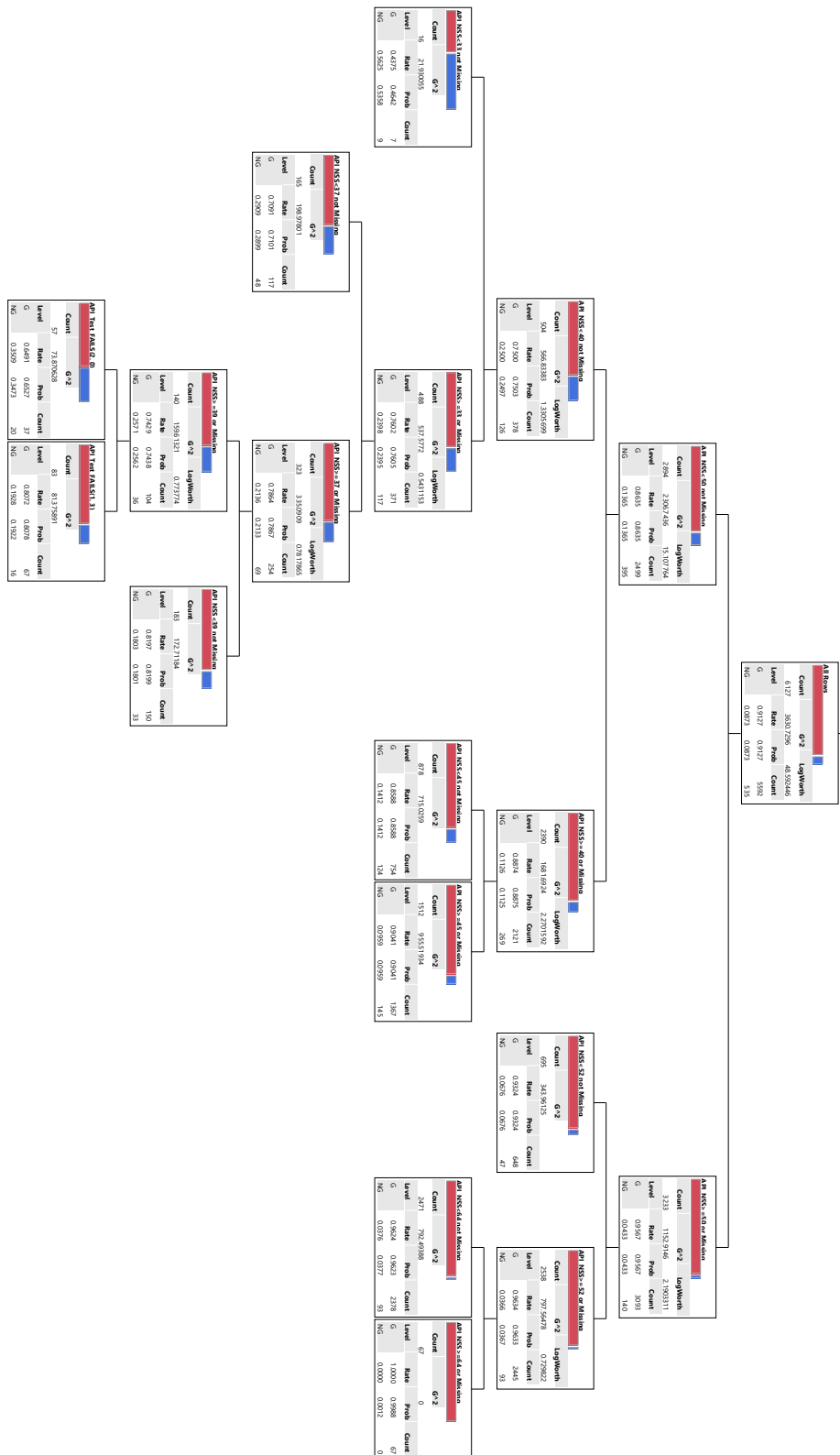


Figure 41. Decision tree model for API

Term	Number of Splits	G ²	Portion
API_NSS	8	270.496116	0.9841
API_Test_FAILS	1	4.3666889	0.0159

Figure 42. Decision tree model column contribution for API

Measure	Training	Validation	Definition
Entropy RSquare	0.0756	0.0680	$1 - \text{Loglike}(\text{model}) / \text{Loglike}(0)$
Generalized RSquare	0.0980	0.0884	$(1 - (L(0)/L(\text{model}))^{(2/n)}) / (1 - L(0)^{(2/n)})$
Mean -Log p	0.2739	0.2767	$\sum -\text{Log}(p[j]) / n$
RMSE	0.2745	0.2757	$\sqrt{\sum (y[j] - \rho[j])^2 / n}$
Mean Abs Dev	0.1507	0.1500	$\sum y[j] - \rho[j] / n$
Misclassification Rate	0.0870	0.0872	$\sum (\rho[j] \neq p\text{Max}) / n$
N	6127	2626	n

Training			Validation		
Actual	Predicted Count		Actual	Predicted Count	
Pri	G	NG	Pri	G	NG
G	5585	7	G	2391	5
NG	526	9	NG	224	6

Figure 43. Decision tree model fit details and confusion matrices for API

Leaf Label	G	NG
API_NSS<50 not Missing&API_NSS<40 not Missing&API_NSS<33 not Missing	0.4642	0.5358
API_NSS<50 not Missing&API_NSS<40 not Missing&API_NSS>=33 or Missing&API_NSS>=37 or Missing&API_NSS>=39 or Missing&API_Test_FAILS(2, 0)	0.6527	0.3473
API_NSS<50 not Missing&API_NSS<40 not Missing&API_NSS>=33 or Missing&API_NSS<37 not Missing	0.7101	0.2899
API_NSS<50 not Missing&API_NSS<40 not Missing&API_NSS>=33 or Missing&API_NSS>=37 or Missing&API_NSS>=39 or Missing&API_Test_FAILS(1, 3)	0.8078	0.1922
API_NSS<50 not Missing&API_NSS<40 not Missing&API_NSS>=33 or Missing&API_NSS>=37 or Missing&API_NSS<39 not Missing	0.8199	0.1801
API_NSS<50 not Missing&API_NSS>=40 or Missing&API_NSS<45 not Missing	0.8588	0.1412
API_NSS<50 not Missing&API_NSS>=40 or Missing&API_NSS>=45 or Missing	0.9041	0.0959
API_NSS>=50 or Missing&API_NSS<52 not Missing	0.9324	0.0676
API_NSS>=50 or Missing&API_NSS>=52 or Missing&API_NSS<64 not Missing	0.9623	0.0377
API_NSS>=50 or Missing&API_NSS>=52 or Missing&API_NSS>=64 or Missing	0.9988	0.0012

Figure 44. Decision tree model leaf report for API

c. Model Comparison

We compare the two developed models to select a best model. Figure 45 and Figure 46 compare these models for both the training and validation data sets. As

shown in these figures, the Decision (Partition) Model has a slightly better R^2 for both the training and validation data sets.

Creator		Entropy RSquare	Generalized RSquare	Mean -Log p	RMSE	Mean Abs Dev	Misclassification Rate	N
Partition		0.0756	0.0980	0.2739	0.2745	0.1507	0.0870	6127
Fit Nominal Logistic		0.0706	0.0917	0.2754	0.2753	0.1519	0.0873	6126

Figure 45. Model comparison for training set for API

Creator		Entropy RSquare	Generalized RSquare	Mean -Log p	RMSE	Mean Abs Dev	Misclassification Rate	N
Partition		0.0680	0.0884	0.2767	0.2757	0.1500	0.0872	2626
Fit Nominal Logistic		0.0673	0.0874	0.2769	0.2757	0.1509	0.0872	2626

Figure 46. Model comparison for validation set for API

d. Conclusions

In this section, we developed a multiple logistics model and a decision tree model to predict the outcome of primary training (PRI) (Graduate or Non-Graduate/Attrite) from API data. Both models developed in this section do not seem to be a good fit as evident by the small R^2 values of the validation set ($<7\%$), which indicate that only a small percentage of the dependent variable variation is explained by the model. Based on the analysis of the developed models, we make the following conclusions:

- The most significant effect in the logistics model for predicting the outcome of primary training is the overall performance of API as Navy Standard Score (API_NSS).
- Similarly, for the decision tree model the most significant variable for predicting the outcome of primary training is the Navy Standard Score (API_NSS).

IV. MODELING ASTBE DATA TO PREDICT PRIMARY TRAINING OUTCOME

In this chapter we use ASTBE data to predict the outcome of Primary training.

1. Predicting Outcome of Primary Training (PRI) from ASTBE Scores

In this section, we develop multiple logistics and bootstrap (random) forests models to predict the outcome of primary training (PRI) (Graduate or Non-Graduate/Attrite) from a set of predictors that include Gender, Race, Hours of previous formal flight instruction, Recruiting Branch, Examinee status, and a set of ASTBE exam scores data. Observations including values other than Graduate (G) or Non-Graduate (NG) in the outcome variable are removed as they are not of interest for this model. We note that, unlike the ASTB345 case, the ASTBE predictors have generally a weak correlation with each other, as shown in Figure 47.

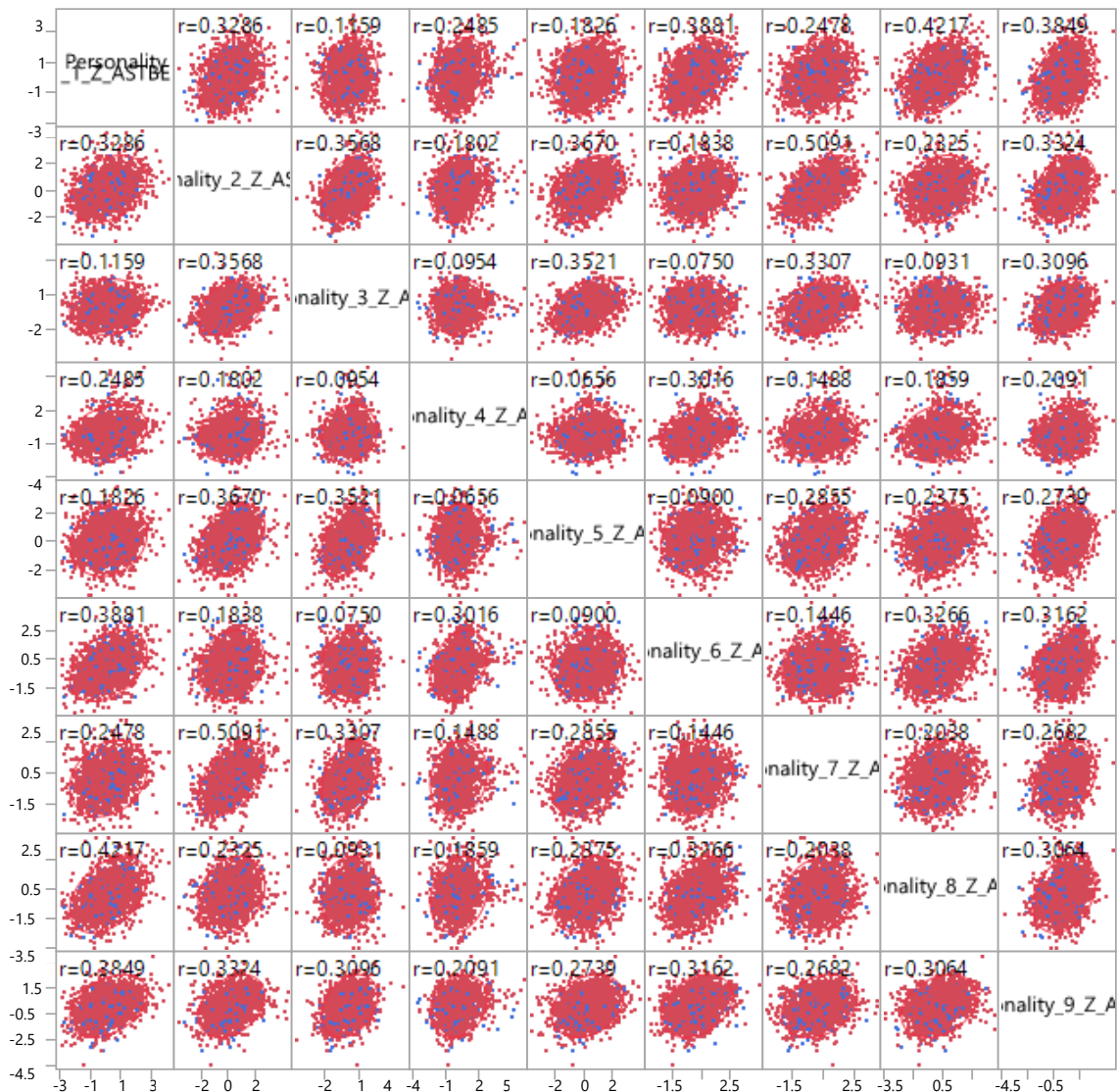


Figure 47. Correlations of ASTBE predictors

a. Multiple Logistics Model with Backward Elimination

This multiple logistics model predicts the outcome of primary training from the full set of potential predictors (main effects). The model is then reduced using backward elimination to remove non-significant predictors. Figure 48 shows the effect summary of the model after removing the insignificant predictors.

Source	LogWorth		PValue
OAR_Z_ASTBE	4.564		0.00003
Personality_9_Z_ASTBE	3.616		0.00024
DLTfactor_Z_ASTBE	1.702		0.01988
Race	1.507		0.03113

Source	LogWorth	PValue
ANIT_Z_ASTBE	1.380	0.04171

Figure 48. Logistics model effect summary for ASTBE

As shown in Figure 48, the top three significant effects for predicting the outcome of primary training are the z-scores for Officer Aptitude Rating on ASTBE (OAR_Z_ASTBE), Personality 9 Rating (Personality_9_Z_ASTBE), and Dichotic Listening Task on ASTBE (DLTFactor_Z_ASTBE) The Whole Model Test Report, shown in Figure 49, has a small p-value (<0.0001), which indicates that the overall model is significant, and an R^2 of 0.0637 indicating that 6.3% of the variation in the response is explained by the model.

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	28.79907	10	57.59813	$<.0001^*$
Full	423.61618			
Reduced	452.41524			
	RSquare (U)		0.0637	
	AICc		869.397	
	BIC		928.511	
	Observations (or Sum Wgts)		1618	

Figure 49. Logistics model whole model test report for ASTBE

The Fit Details Report, shown in Figure 50, indicates the model has a misclassification rate of about 8% for both the training and validation sets. However, this low misclassification rate is misleading due to the data imbalance in the response variable, since the number of Non-Graduates is only about 8% vs. 92% for Graduates.

Measure	Training	Validation	Definition
Entropy RSquare	0.0637	-0.086	$1 - \text{Loglike}(\text{model}) / \text{Loglike}(0)$
Generalized RSquare	0.0816	-0.114	$(1 - (L(0)/L(\text{model}))^{(2/n)}) / (1 - L(0)^{(2/n)})$
Mean -Log p	0.2618	0.3009	$\sum -\text{Log}(p[j]) / n$
RMSE	0.2664	0.2781	$\sqrt{\sum (y[j] - p[j])^2 / n}$
Mean Abs Dev	0.1416	0.1486	$\sum y[j] - p[j] / n$
Misclassification Rate	0.0803	0.0851	$\sum (p[j] \neq pMax) / n$
N	1618	693	n

Figure 50. Logistics model fit details report for ASTBE

Figure 51 shows the confusion matrix for both the training and validation sets.

Training			Validation		
Actual	Predicted Count		Actual	Predicted Count	
Pri	NG	G	Pri	NG	G
NG	1	129	NG	0	55
G	1	1487	G	4	634

Figure 51. Logistics model confusion matrix for ASTBE

As a result of the data imbalance, the model does not predict any Non-Graduates. To remedy this situation and specify a better classifier, a cut off lower than 0.5 needs to be specified for classifying a predicted probability as Non-Graduates.

Figure 52 shows the prediction profiler. It indicates the predicted probability of the response variable PRI for each specified value of the predictor variables. As shown in the figure, a higher z-score for Officer Aptitude Rating on ASTBE (OAR_Z_ASTBE), Personality 9 Rating (Personality_9_Z_ASTBE), Dichotic Listening Task on ASTBE (DLTFactor_Z_ASTBE), and Aviation Nautical Info on ASTBE (ANIT_Z_ASTBE) is associate with a lower probability of Non-Graduate outcome. We note that the model is unstable for Race due to insufficient data, and therefore no conclusions can be drawn from the race variable.

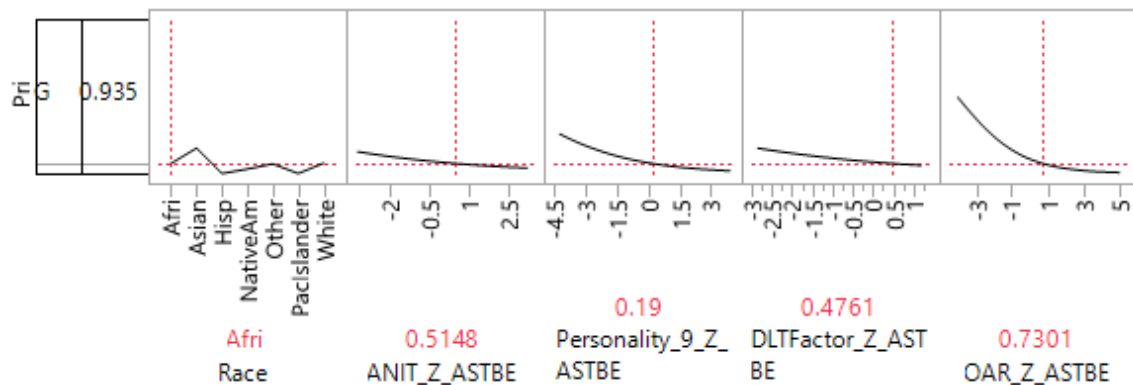


Figure 52. Logistics model prediction profiler for ASTBE

b. Bootstrap (Random) Forest Model

This model uses a bootstrap (random) forest to predict the outcome of primary training from the full set of potential predictors. Figure 53 shows the overall

model summary statistics. For the validation set, R^2 is 0.0024 and the misclassification rate is 0.0794.

Measure	Training	Validation	Definition
Entropy RSquare	0.4485	0.0024	$1 - \text{Loglike}(\text{model}) / \text{Loglike}(0)$
Generalized RSquare	0.5179	0.0031	$(1 - (L(0)/L(\text{model}))^{(2/n)}) / (1 - L(0)^{(2/n)})$
Mean -Log p	0.1542	0.2765	$\sum -\text{Log}(p[j]) / n$
RMSE	0.2103	0.2708	$\sqrt{\sum (y[j] - p[j])^2 / n}$
Mean Abs Dev	0.1089	0.1420	$\sum y[j] - p[j] / n$
Misclassification Rate	0.0803	0.0794	$\sum (p[j] \neq p_{\text{Max}}) / n$
N	1618	693	n

Figure 53. Bootstrap forest overall model summary statistics for ASTBE

Figure 54 shows the column contribution of the different predictors with the z-scores for Dichotic Listening Task (DLTFactor_Z_ASTBE), Flight Officer Aptitude Rating (FOFAR_Z_ASTBE), and Personality 9 Rating (Personality_9_Z_ASTBE) being the top three predictor contributors.

Term	Number of Splits	G ²	Portion
DLTFactor_Z_ASTBE	155	17.7184526	0.0598
FOFAR_Z_ASTBE	113	16.6782539	0.0562
Personality_9_Z_ASTBE	128	16.6076618	0.0560
AQR_Z_ASTBE	119	15.6606621	0.0528
Personality_7_Z_ASTBE	131	14.799905	0.0499
SkillFactor_Z_ASTBE	126	13.7377453	0.0463
RCT_Z_ASTBE	113	13.6867111	0.0462
MCT_Z_ASTBE	123	13.4935971	0.0455
ANIT_Z_ASTBE	110	12.3780123	0.0417
Personality_4_Z_ASTBE	111	12.2416541	0.0413
Personality_3_Z_ASTBE	113	11.8535405	0.0400
OAR_Z_ASTBE	92	11.5795418	0.0390
Personality_2_Z_ASTBE	115	11.126458	0.0375
VTTFactor_Z_ASTBE	104	11.0449332	0.0372
Personality_5_Z_ASTBE	104	10.9856162	0.0370
Personality_6_Z_ASTBE	104	10.8916103	0.0367
DOTFactor_Z_ASTBE	99	10.7400198	0.0362
MST_Z_ASTBE	108	10.4304008	0.0352
Personality_2_Z_ASTBE	99	10.2134603	0.0344
Personality_8_Z_ASTBE	86	9.08366318	0.0306
ATTFactor_Z_ASTBE	84	8.98712963	0.0303
PFAR_Z_ASTBE	91	8.60774856	0.0290
Race	68	7.92964669	0.0267
FormalFlightInstrHours	80	6.27743036	0.0212
ExamineeStatus_A	96	5.5987421	0.0189
Gender	61	2.50293908	0.0084
RecruitingBranch_A	41	1.68123387	0.0057

Figure 54. Bootstrap forest model column contributions for ASTBE

Figure 55 shows the confusion matrix for both the training and validation sets.

Training			Validation		
Actual	Predicted Count		Actual	Predicted Count	
Pri	G	NG	Pri	G	NG
G	1488	0	G	638	0
NG	130	0	NG	55	0

Figure 55. Bootstrap forest model confusion matrix for ASTBE

As indicated earlier and as a result of the data imbalance, the model does not predict any Non-Graduates. Similar to the approach used in the logistics regression model, we need to specify a probability cut off lower than 0.5 to classify more observations as Non-Graduates.

Figure 56 shows the prediction profiler. It indicates the predicted probability of the response variable PRI for each specified value of the predictor variables. As shown in the figure, there is very little or no effect of the predictor variables on the target variable across all ranges of predictor values.

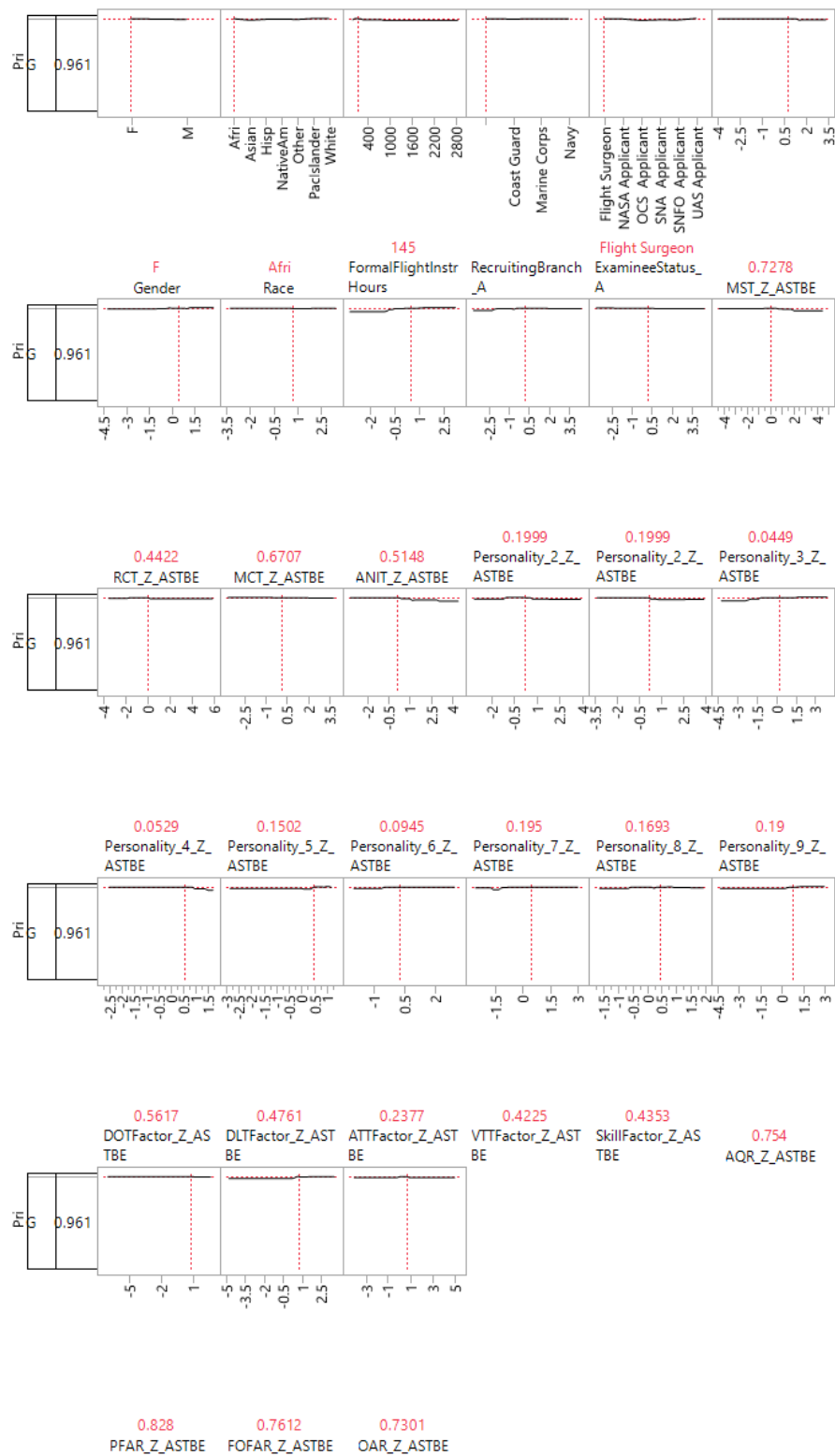


Figure 56. Bootstrap forest model prediction profiler for ASTBE

c. Model Comparison

We compare the developed models to select a best model. Figure 57 and Figure 58 compare these models for both the training and validation data sets. As shown in these figures, the Bootstrap Forest Model has a better R^2 for both the training and validation data sets

Creator		Entropy RSquare	Generalized RSquare	Mean - Log p	RMSE	Mean Abs Dev	Misclassification Rate	N
Bootstrap Forest		0.4485	0.5179	0.1542	0.2103	0.1089	0.0803	1618
Fit Nominal Logistic		0.0637	0.0816	0.2618	0.2664	0.1416	0.0803	1618

Figure 57. Model comparison for training set for ASTBE

Creator		Entropy RSquare	Generalized RSquare	Mean - Log p	RMSE	Mean Abs Dev	Misclassification Rate	N
Bootstrap Forest		0.0024	0.0031	0.2765	0.2708	0.1420	0.0794	693
Fit Nominal Logistic		-0.086	-0.114	0.3009	0.2781	0.1486	0.0851	693

Figure 58. Model comparison for validation set for ASTBE

d. Conclusions

In this section, we developed a multiple logistics model and a bootstrap (random) forest model to predict the outcome of primary training (PRI) (Graduate or Non-Graduate/Attrite) from ASTBE data. Both models developed in this section do not seem to be a good fit as evident by the small R^2 values of the validations set, which indicate that only a small percentage of the dependent variable variation is explained by the model. Based on the analysis of the developed models, we make the following conclusions:

- For the multiple logistics model, the top three significant effects for predicting the outcome of primary training are the z-scores for Officer Aptitude Rating on ASTBE (OAR_Z_ASTBE), Personality 9 Rating (Personality_9_Z_ASTBE), and Dichotic Listening Task on ASTBE (DLTFactor_Z_ASTBE).
- For bootstrap forest model, the top three variable contributors are the z-scores for Dichotic Listening Task (DLTFactor_Z_ASTBE), Flight Officer Aptitude Rating (FOFAR_Z_ASTBE), and Personality 9 Rating (Personality_9_Z_ASTBE).

2. Predicting Outcome of Primary Training (PRI) from IFS Scores

In this section, we develop multiple logistics models and a bootstrap (random) forest model to predict the outcome of primary training (PRI) (Graduate or Non-Graduate/Attrite) from IFS data for students who took the ASTBE. IFS Predictors include number of hours of flight time received during IFS, number of landings completed in IFS, exam score on Stages 1, 2, and 3 Exams, End of Course exam score, FAA exam score, and number of flight failures in IFS. We note that there exists some correlation between the IFS exam scores, as shown in Figure 59.

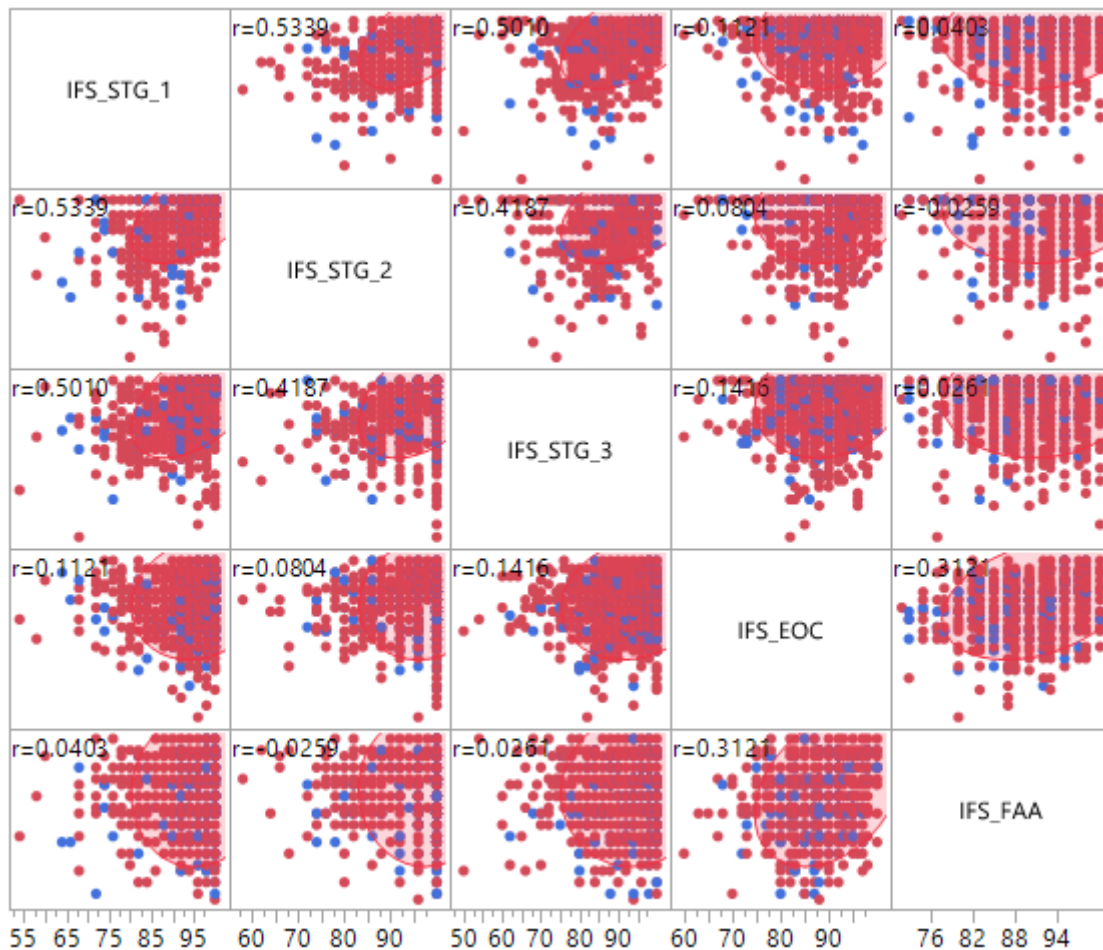


Figure 59. Correlation of IFS scores

a. Multiple Logistics Model with Backward Elimination

This multiple logistics model predicts the outcome of primary training from the full set of potential predictors (main effects). The model is then reduced using

backward elimination to remove non-significant predictors. Figure 60 shows the effect summary of the model after removing the insignificant predictors.

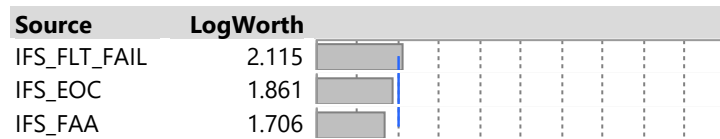


Figure 60. Logistics model effect summary for IFS

As shown in Figure 60, the top three significant effects for predicting the outcome of primary training are the number of flight failures in IFS, the End of Course Exam Score, and the FAA Exam Score. The Whole Model Test Report, shown in Figure 61, has a small p-value (<0.0001), which indicates that the overall model is significant, and an R^2 of 0.0474 indicating that 4.74 % of the variation in response is explained by the model.

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	14.45731	4	28.91462	<.0001*
Full	290.75997			
Reduced	305.21727			
	RSquare (U)		0.0474	
	AICc		591.576	
	BIC		616.457	
	Observations (or Sum Wgts)		1083	

Figure 61. Logistics model whole model test for IFS

Figure 62 shows the confusion matrix for both the training and validation sets.

Training			Validation		
Actual	Predicted Count		Actual	Predicted Count	
Pri	NG	G	Pri	NG	G
NG	2	86	NG	0	41
G	2	993	G	2	418

Figure 62. Logistics model confusion matrix for IFS

As a result of the data imbalance, the model performs poorly for predicting Non-Graduates. To remedy this situation, and as discussed earlier, a cut off

lower than 0.5 needs to be specified for classifying a predicted probability as Non-Graduates.

Figure 63 shows the prediction profiler. It indicates the predicted probability of the response variable PRI for each specified value of the predictor variables. As shown in the figure, a higher IFS End of Course Score and IFS FAA Score, indicate a lower probability of Non-Graduate outcome, while a higher number of IFS flight failures is associated with a higher probability of Non-Graduate outcome in Primary training.

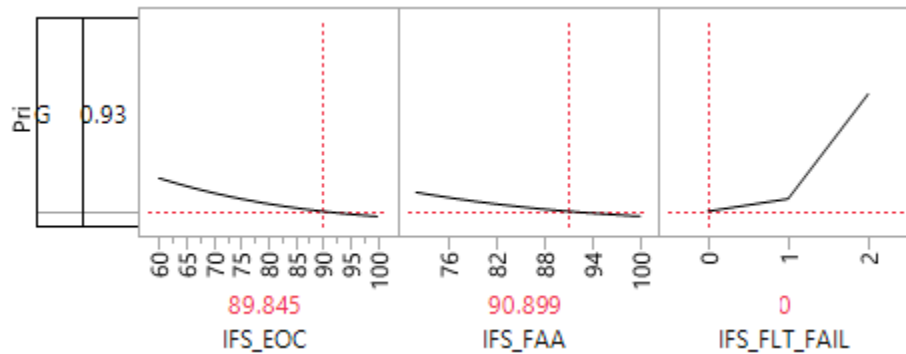


Figure 63. Logistics model prediction profiler for IFS

b. Bootstrap (Random) Forest Model

This model uses a bootstrap (random) forest to predict the outcome of primary training from the full set of potential predictors (main effects). Figure 64 shows the overall model summary statistics. For the validation set, R^2 is 0.0097 and the misclassification rate is 0.0794.

Measure	Training	Validation	Definition
Entropy RSquare	0.1556	0.0097	$1 - \text{Loglike}(\text{model}) / \text{Loglike}(0)$
Generalized RSquare	0.1945	0.0127	$(1 - (L(0) / L(\text{model}))^{(2/n)}) / (1 - L(0)^{(2/n)})$
Mean -Log p	0.2361	0.2745	$\sum -\text{Log}(p[j]) / n$
RMSE	0.2557	0.2696	$\sqrt{\sum (y[j] - p[j])^2 / n}$
Mean Abs Dev	0.1371	0.1489	$\sum y[j] - p[j] / n$
Misclassification Rate	0.0803	0.0794	$\sum (p[j] \neq p\text{Max}) / n$
N	1618	693	n

Figure 64. Bootstrap forest overall model summary for IFS

Figure 65 shows the column contribution of the different predictors with number of landings completed during IFS (IFS_TOTAL_LANDINGS), End of Course

Exam Score (IFS_EOC), and FAA Exam Score (IFS_FAA), being the top three contributors.

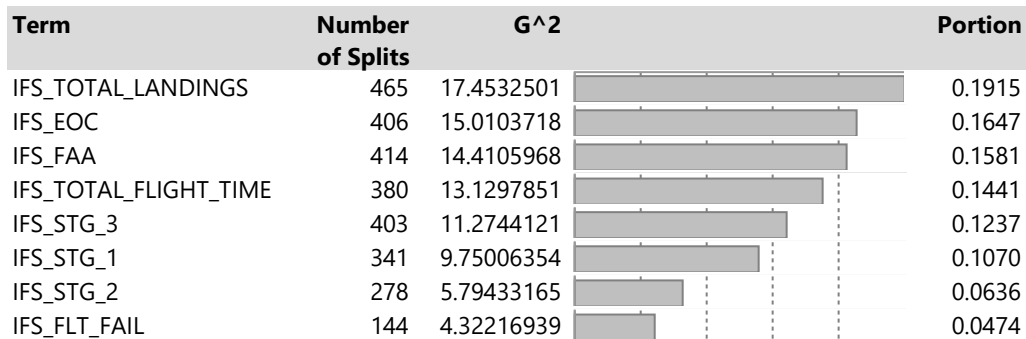


Figure 65. Bootstrap forest model column contributions for IFS

Figure 66 shows the confusion matrix for both the training and validation sets.

Training			Validation		
Actual	Predicted Count		Actual	Predicted Count	
Pri	G	NG	Pri	G	NG
G	1488	0	G	638	0
NG	130	0	NG	55	0

Figure 66. Bootstrap forest model confusion matrix for IFS

As indicated earlier and as a result of the data imbalance, although the misclassification rate is low, the model does not predict any Non-Graduates. Similar to the approach used in the logistics regression model, we need to specify a probability cut off lower than 0.5 to classify more observations as Non-Graduates.

Figure 67 shows the prediction profiler. It indicates the predicted probability of the response variable PRI for each specified value of the predictor variables. As shown in the Figure 67, the response variable PRI has a low sensitivity with respect to the predictor variables.

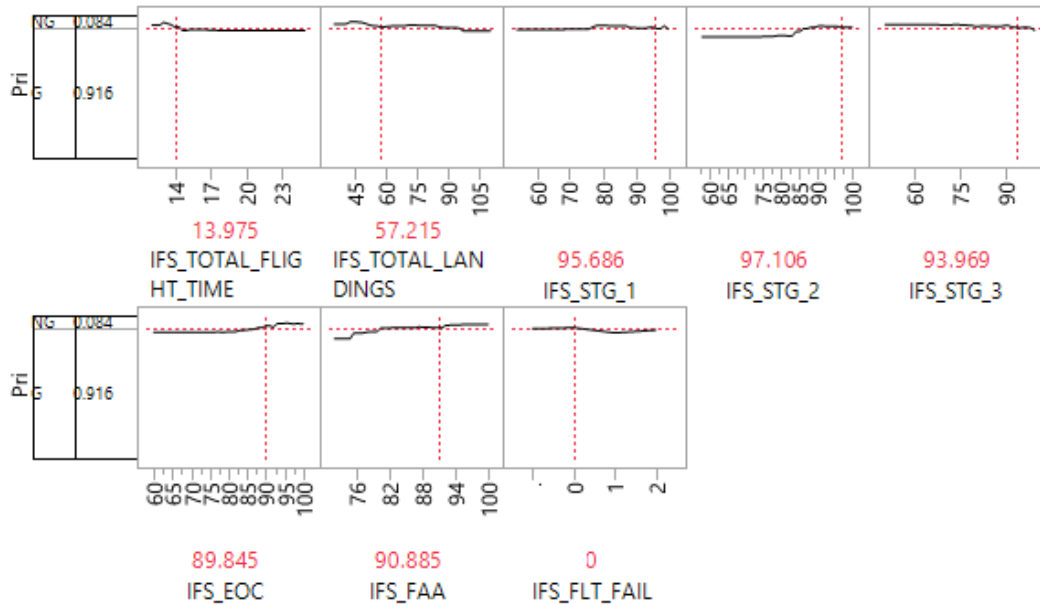


Figure 67. Bootstrap forest model prediction profiler for IFS

c. Model Comparison

We compare the developed models to select a best model. Figure 68 and Figure 69 compare these models for both the training and validation data sets. As shown in these figures, the Bootstrap Forest Model has a better R^2 for both the training and validation data sets

Creator		Entropy RSquare	Generalized RSquare	Mean - Log p	RMSE	Mean Abs Dev	Misclassification Rate	N
Bootstrap Forest		0.1556	0.1945	0.2361	0.2557	0.1371	0.0803	1618
Fit Nominal Logistic		0.0474	0.0611	0.2685	0.2687	0.1441	0.0813	1083

Figure 68. Model comparison for training set for IFS

Creator		Entropy RSquare	Generalized RSquare	Mean - Log p	RMSE	Mean Abs Dev	Misclassification Rate	N
Bootstrap Forest		0.0097	0.0127	0.2745	0.2696	0.1489	0.0794	693
Fit Nominal Logistic		0.0043	0.0058	0.2988	0.2863	0.1613	0.0933	461

Figure 69. Model comparison for validation set for IFS

d. Conclusions

In this section, we developed a multiple logistics model and a bootstrap (random) forest model to predict the outcome of primary training (PRI) (Graduate or Non-

Graduate/Attrite) from IFS data. Both models developed in this section do not seem to be a good fit as evident by the small R^2 values, which indicate that only a small percentage of the dependent variable variation is explained by the model. Based on the analysis of the developed models, we make the following conclusions:

- The top three significant effects in the logistics model for predicting the outcome of primary training are the number of flight failures in IFS (IFS_FLT_FAIL), the End of Course Exam Score (IFS_EOC), and the FAA Exam Score (IFS_FAA).
- For bootstrap forest model, the top three variable contributors are number of landings completed during IFS (IFS_TOTAL_LANDINGS), End of Course Exam Score (IFS_EOC), and FAA Exam Score (IFS_FAA).
- The bootstrap forest model performed better on both the training and validation data

3. Predicting Outcome of Primary Training (PRI) from API Scores

In this section, we develop a multiple logistics model and a decision (partition) tree models to predict the outcome of primary training (PRI) (Graduate or Non-Graduate/Attrite) from API data for students who took the ASTBE. API Predictors include the overall performance of API as Navy Standard Score and the number of academic failures in API.

a. Multiple Logistics Model

This multiple logistics model predicts the outcome of primary training from the set of potential predictors (main effects). Figure 70 shows the effect summary of the model.

Source	LogWorth	PValue
API_NSS	7.409	0.00000
API_Test_FAILS	0.003	0.99279

Figure 70. Logistics model effect summary for API

As shown in Figure 70, the most significant effect for predicting the outcome of primary training is the overall performance of API as Navy Standard Score (API_NSS). The Whole Model Test Report, shown in Figure 71, has a small p-value

(<0.0001), which indicates that the overall model is significant, and an R^2 of 0.0589 indicating that about 6 % of the variation in response is explained by the model.

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	26.65112	2	53.30224	<.0001*
Full	425.59650			
Reduced	452.24762			
RSquare (U)			0.0589	
AICc			857.208	
BIC			873.356	
Observations (or Sum Wgts)			1616	

Figure 71. Logistics model whole model test for API

Figure 72 shows the confusion matrix for both the training and validation sets.

Training			Validation		
Actual	Predicted Count		Actual	Predicted Count	
Pri	NG	G	Pri	NG	G
NG	0	130	NG	0	55
G	0	1486	G	0	637

Figure 72. Logistics model confusion matrix for API

As discussed earlier and as a result of the data imbalance, the model does not predict any Non-Graduates. To remedy this situation, we need to specify a cut off lower than 0.5 for classifying a predicted probability as Non-Graduates.

Figure 73 shows the prediction profiler. It indicates the predicted probability of the response variable PRI for each value of the API predictor variables. As shown in the figure, a higher value of the overall performance of API as Navy Standard score (API_NSS) is associated with a lower probability of Non-Graduate outcome, while a higher number of academic failures in API (API_Test_FAILS) is associated with a higher probability of Non-Graduate.

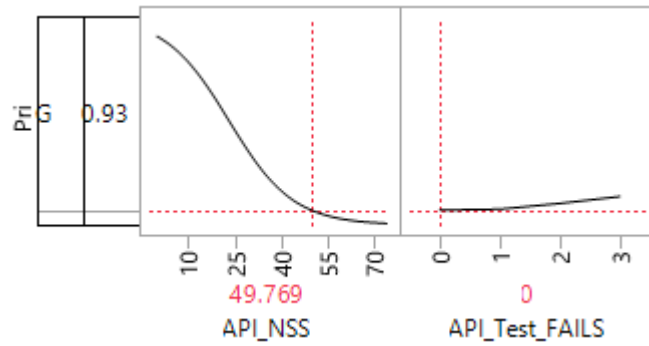


Figure 73. Logistics model prediction profiler for API

b. Decision (Partition) Tree

In this section, we fit a decision tree model to the API data for students who took the ASTBE Exam. Figure 74 shows the resulting decision tree pruned using the validation data to avoid overfitting. Most of the splits occur on the API Navy Standard Score (API_NSS), which is the most significant predictor, as shown in Figure 75. The fit details and confusion matrices are shown in

Figure 76, and the leaf report, shown in Figure 77, displays the rules generated by the decision tree. As indicated by the first rule of the figure, a Navy Standard Score of less than 34 is associated with a high probability of Non-Graduate Outcome.

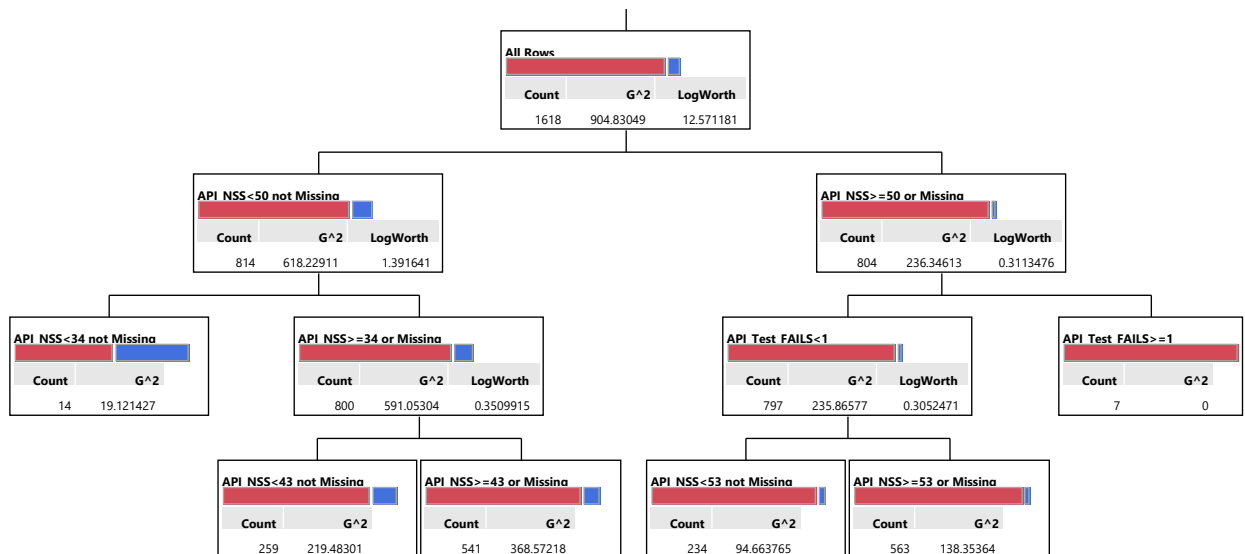


Figure 74. Decision tree for API

Term	Number of Splits	G ²	Portion
API_NSS	4	64.1561078	0.9926
API_Test_FAILS	1	0.48035524	0.0074

Figure 75. Decision tree model column contribution for API

Measure	Training	Validation	Definition
Entropy RSquare	0.0713	0.0612	$1 - \text{Loglike}(\text{model}) / \text{Loglike}(0)$
Generalized RSquare	0.0912	0.0784	$(1 - (L(0)/L(\text{model}))^{(2/n)}) / (1 - L(0)^{(2/n)})$
Mean -Log p	0.2597	0.2602	$\sum -\text{Log}(p[j]) / n$
RMSE	0.2658	0.2666	$\sqrt{\sum (y[j] - \rho[j])^2 / n}$
Mean Abs Dev	0.1413	0.1438	$\sum y[j] - \rho[j] / n$
Misclassification Rate	0.0803	0.0794	$\sum (p[j] \neq p_{\text{Max}}) / n$
N	1618	693	n

Training			Validation		
Actual	Predicted Count		Actual	Predicted Count	
Pri	G	NG	Pri	G	NG
G	1488	0	G	638	0
NG	130	0	NG	55	0

Figure 76. Decision tree model fit details and confusion matrices for API

Leaf Label	G	NG
API_NSS<50 not Missing&API_NSS<34 not Missing	0.5943	0.4057
API_NSS<50 not Missing&API_NSS>=34 or Missing&API_NSS<43 not Missing	0.8497	0.1503
API_NSS<50 not Missing&API_NSS>=34 or Missing&API_NSS>=43 or Missing	0.8928	0.1072
API_NSS>=50 or Missing&API_Test_FAILS<1&API_NSS<53 not Missing	0.9486	0.0514
API_NSS>=50 or Missing&API_Test_FAILS<1&API_NSS>=53 or Missing	0.9733	0.0267
API_NSS>=50 or Missing&API_Test_FAILS>=1	0.9905	0.0095

Figure 77. Decision tree model leaf report for API

c. Model Comparison

We compare the two developed models to select a best model. Figure 78 and Figure 79 compare these models for both the training and validation data sets. As shown in these figures, the Decision (Partition) Model has a better R^2 for the training data, while the Logistic model is the better model for the validation data set.

Creator	Entropy RSquare	Generalized RSquare	Mean - Log p	RMSE	Mean Abs Dev	Misclassification Rate	N
Partition	0.0713	0.0912	0.2597	0.2658	0.1413	0.0803	1618
Fit Nominal Logistic	0.0589	0.0757	0.2634	0.2673	0.1427	0.0804	1616

Figure 78. Model comparison for training set for API

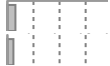
Creator		Entropy RSquare	Generalized RSquare	Mean - Log p	RMSE	Mean Abs Dev	Misclassification Rate	N
Fit Nominal Logistic		0.0729	0.0931	0.2573	0.2648	0.1451	0.0795	692
Partition		0.0612	0.0784	0.2602	0.2666	0.1438	0.0794	693

Figure 79. Model comparison for validation set for API

d. Conclusions

In this section, we developed a multiple logistics model and a decision tree model to predict the outcome of primary training (PRI) (Graduate or Non-Graduate/Attrite) from API data. Both models developed in this section do not seem to be a good fit as evident by the small R^2 values of the validation set (<8%), which indicate that only a small percentage of the dependent variable variation is explained by the model. Based on the analysis of the developed models, we make the following conclusions:

- The most significant effect in the logistics model for predicting the outcome of primary training is the overall performance of API as Navy Standard Score (API_NSS).
- Similarly, most of the splits for the decision tree model occur on the API Navy Standard Score (API_NSS), indicating it is the most significant predictor.
- The decision tree (partition) model is the better model (higher R^2) for both the training and validation data sets.

V. MODELING ASTBE DATA TO PREDICT ADVANCED TRAINING OUTCOME

In this chapter we use ASTBE data to predict the outcome of Advanced training.

1. Predicting Outcome of Advanced Training (ADV) from ASTBE Scores

In this section, we develop multiple logistics and bootstrap (random) forest models to predict the outcome of advanced training (ADV) (Graduate or Non-Graduate/Attrite) from a set of predictors that include Gender, Race, Hours of previous formal flight instruction, Recruiting Branch, Examinee status, and a set of ASTBE exam scores data. Observations including values other than Graduate (G) or Non-Graduate (NG) are removed as they are not of interest for this model.

a. Multiple Logistics Model with Backward Elimination

This multiple logistics model predicts the outcome of advanced training from the full set of potential predictors (main effects). The model is then reduced using backward elimination to remove non-significant predictors. Figure 80 shows the effect summary of the model after removing the insignificant predictors.

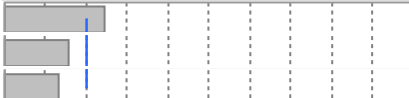
Source	LogWorth		PValue
AQR_Z_ASTBE	2.481		0.00330
OAR_Z_ASTBE	1.602		0.02498
Race	1.347		0.04500

Figure 80. Logistics model effect summary for ASTBE

As shown in Figure 80, the top two significant effects for predicting the outcome of primary training are the z-scores for Academic Qualifications Rating on ASTBE (AQR_Z_ASTBE) and Officer Aptitude Rating on ASTBE (OAR_Z_ASTBE). The Whole Model Test Report, shown in Figure 81, has a small p-value (<0.0001), which indicates that the overall model is significant, and an R^2 of 0.0494 indicating that 4.94 % of the variation in the response is explained by the model.

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	10.37881	7	20.75762	0.0041*
Full	199.66502			

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Reduced	210.04383			
	RSquare (U)		0.0494	
	AICc		415.506	
	BIC		453.072	
	Observations (or Sum Wgts)		827	

Figure 81. Logistics model whole model test report for ASTBE

The Fit Details Report, shown in Figure 82, indicates the model has a misclassification rate of about 7% for both the training and validation sets. However, this low misclassification rate is misleading due to the data imbalance in the response variable, since the number of Non-Graduates is only about 7% vs. 93% for Graduates.

Measure	Training	Validation	Definition
Entropy RSquare	0.0494	0.0475	$1 - \text{Loglike}(\text{model}) / \text{Loglike}(0)$
Generalized RSquare	0.0622	0.0599	$(1 - (L(0)/L(\text{model}))^{(2/n)}) / (1 - L(0)^{(2/n)})$
Mean -Log p	0.2414	0.2431	$\sum -\text{Log}(p[j]) / n$
RMSE	0.2518	0.2509	$\sqrt{\sum (y[j] - p[j])^2 / n}$
Mean Abs Dev	0.1267	0.1259	$\sum y[j] - p[j] / n$
Misclassification Rate	0.0701	0.0678	$\sum (p[j] \neq p\text{Max}) / n$
N	827	354	n

Figure 82. Logistics model fit details report for ASTBE

Figure 83 shows the confusion matrix for both the training and validation sets.

Training			Validation		
Actual	Predicted Count		Actual	Predicted Count	
Pri	NG	G	Pri	NG	G
NG	0	58	NG	1	24
G	0	769	G	0	329

Figure 83. Logistics model confusion matrix for ASTBE

As a result of the data imbalance, the model behaves very poorly for predicting Non-Graduates. To remedy this situation and specify a better classifier, a cut off lower than 0.5 needs to be specified for classifying a predicted probability as Non-Graduates.

Figure 84 shows the prediction profiler. It indicates the predicted probability of the response variable PRI for each specified value of the predictor

variables. As shown in the figure, for this reduced model, a higher z-score for Academic Qualifications Rating on ASTBE (AQR_Z_ASTBE) is associated with a lower probability of Non-Graduate outcome, while a higher z-score for Officer Aptitude Rating on ASTBE (OAR_Z_ASTBE) is associated with a higher probability of Non-Graduate outcome. We note that the model is unstable for Race due to insufficient data, and therefore no conclusions can be drawn from the Race variable.

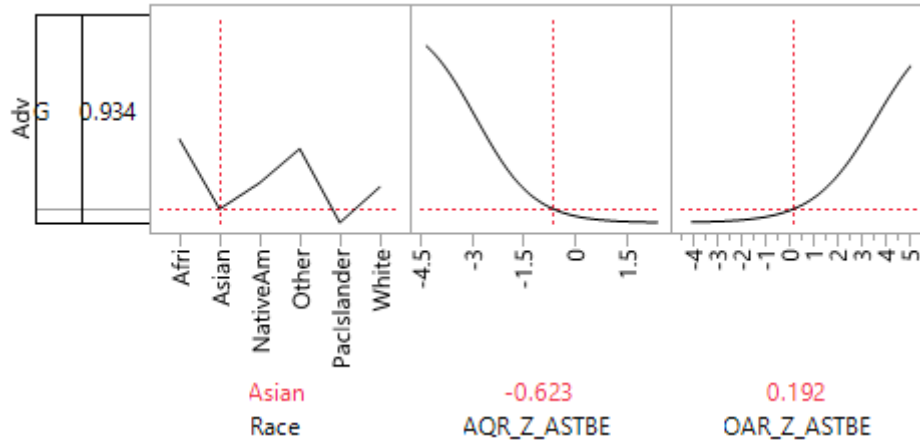


Figure 84. Logistics model prediction profiler for ASTBE

b. Bootstrap (Random) Forest Model

This model uses a bootstrap (random) forest to predict the outcome of primary training from the full set of potential predictors. Figure 85 shows the overall model summary statistics. For the validation set, R^2 is 0.013 and the misclassification rate is 0.0706.

Measure	Training	Validation	Definition
Entropy RSquare	0.4414	-0.013	$1 - \text{Loglike}(\text{model}) / \text{Loglike}(0)$
Generalized RSquare	0.5044	-0.017	$(1 - (L(0)/L(\text{model}))^{(2/n)}) / (1 - L(0)^{(2/n)})$
Mean -Log p	0.1419	0.2586	$\sum -\text{Log}(p[j]) / n$
RMSE	0.2007	0.2572	$\sqrt{\sum (y[j] - p[j])^2 / n}$
Mean Abs Dev	0.0985	0.1256	$\sum y[j] - p[j] / n$
Misclassification Rate	0.0701	0.0706	$\sum (p[j] \neq p\text{Max}) / n$
N	827	354	n

Figure 85. Bootstrap forest overall model summary statistics for ASTBE

Figure 86 shows the column contribution of the different predictors with the z-scores for Personality 9 Rating (Personality_9_Z_ASTBE), Academic

Qualifications Rating on ASTBE (AQR_Z_ASTBE), and Dichotic Listening Task (DLTFactor_Z_ASTBE) being the top three predictor contributors.

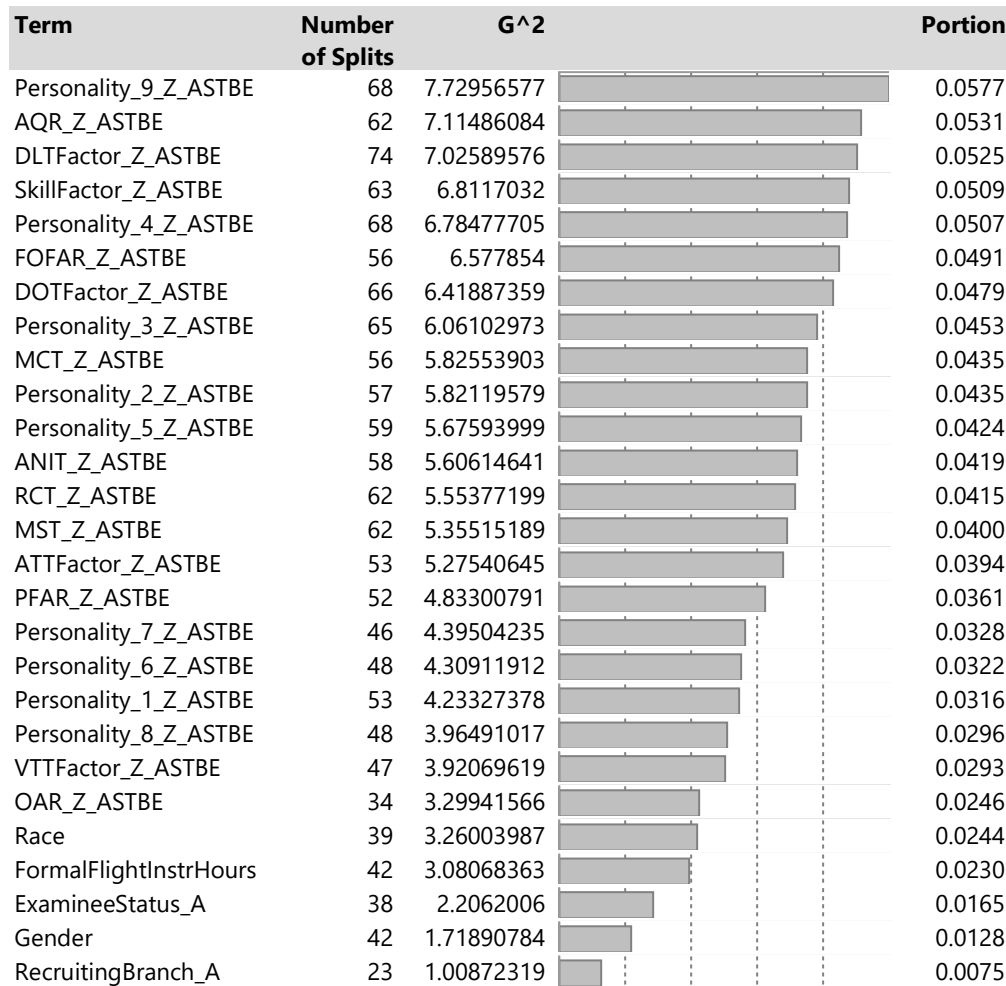


Figure 86. Bootstrap forest model column contribution for ASTBE

Figure 87 shows the confusion matrix for both the training and validation sets.

Training			Validation		
Actual	Predicted Count		Actual	Predicted Count	
Pri	G	NG	Pri	G	NG
G	769	0	G	329	0
NG	58	0	NG	25	0

Figure 87: Bootstrap forest model confusion matrix for ASTBE

As indicated earlier and as a result of the data imbalance, the model does not predict any Non-Graduates. Similar to the approach used in the logistics regression model, we need to specify a probability cut off lower than 0.5 to classify more observations as Non-Graduates.

Figure 88 shows the prediction profiler. It indicates the predicted probability of the response variable ADV for each specified value of the predictor variables. As shown in the figure, there is very little or no effect of the predictor variables on the target variable across all ranges of predictor values.

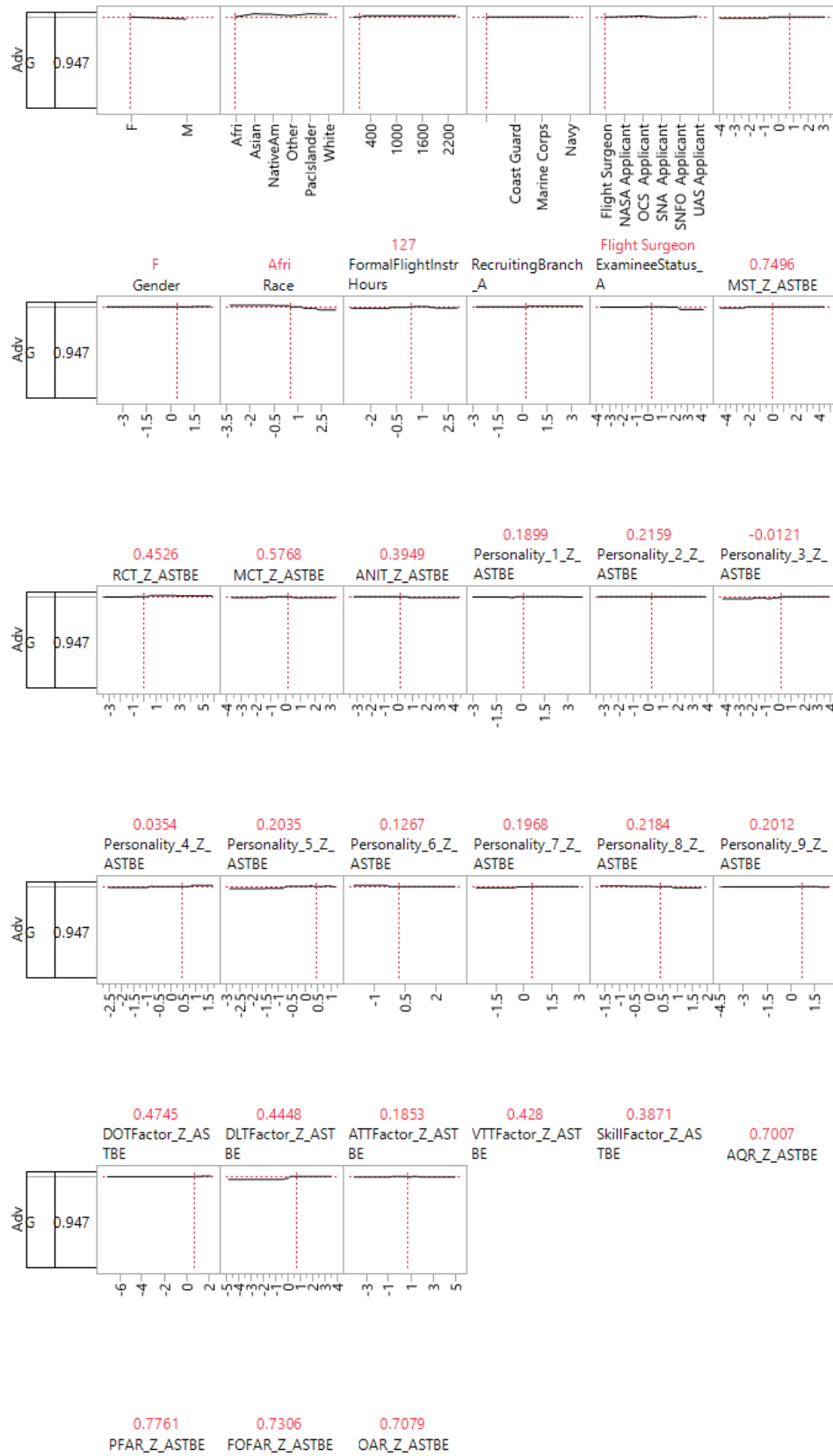


Figure 88. Bootstrap forest model prediction profiler for ASTBE

c. Model Comparison

We compare the developed models to select a best model. Figure 89 and Figure 90 compare these models for both the training and validation data sets. As shown in these figures, the Bootstrap Forest Model has a better R^2 for the training data set, while the Nominal Logistics Model has a better R^2 for the validation data set.

Creator		Entropy RSquare	Generalized RSquare	Mean - Log p	RMSE	Mean Abs Dev	Misclassification Rate	N
Bootstrap Forest		0.4414	0.5044	0.1419	0.2007	0.0985	0.0701	827
Fit Nominal Logistic		0.0494	0.0622	0.2414	0.2518	0.1267	0.0701	827

Figure 89. Model comparison for training set for ASTBE

Creator		Entropy RSquare	Generalized RSquare	Mean - Log p	RMSE	Mean Abs Dev	Misclassification Rate	N
Fit Nominal Logistic		0.0475	0.0599	0.2431	0.2509	0.1259	0.0678	354
Bootstrap Forest		-0.013	-0.017	0.2586	0.2572	0.1256	0.0706	354

Figure 90. Model comparison for validation set for ASTBE

d. Conclusions

In this section, we developed a multiple logistics model and a bootstrap (random) forest model to predict the outcome of advanced training (ADV) (Graduate or Non-Graduate/Attrite) from ASTBE data. Both models developed in this section do not seem to be a good fit as evident by the small R^2 values of the validation set, which indicate that only a small percentage of the dependent variable variation is explained by the model. Based on the analysis of the developed models, we make the following conclusions:

- For the multiple logistics model, the top two significant effects for predicting the outcome of advanced training are the z-scores for Academic Qualifications Rating on ASTBE (AQR_Z_ASTBE) and Officer Aptitude Rating on ASTBE (OAR_Z_ASTBE).
- For bootstrap forest model, the top three variable contributors are the z-scores for the Personality 9 Rating (Personality_9_Z_ASTBE), Academic Qualifications Rating on ASTBE (AQR_Z_ASTBE), and Dichotic Listening Task (DLTFactor_Z_ASTBE).

- The Bootstrap Forest Model has a better R^2 for the training data set, while the Nominal Logistics Model has a better R^2 for the validation data set.

2. Predicting Outcome of Advanced Training (ADV) from IFS Scores

In this section, we develop multiple logistics models and a bootstrap (random) forest model to predict the outcome of advanced training (ADV) (Graduate or Non-Graduate/Attrite) from IFS data for students who took the ASTBE. IFS Predictors include number of hours of flight time received during IFS, number of landings completed in IFS, exam score on Stages 1, 2, and 3 Exams, End of Course exam score, FAA exam score, and number of flight failures in IFS.

a. Multiple Logistics Model with Backward Elimination

This multiple logistics model predicts the outcome of advanced training from the full set of potential predictors (main effects). The model is then reduced using backward elimination to remove non-significant predictors. Figure 91 shows the effect summary of the model after removing the insignificant predictors.

Source	LogWorth		PValue
IFS_FAA	2.489		0.00324
IFS_STG_1	2.028		0.00938

Figure 91: Logistics model effect summary for IFS

As shown in Figure 91, the two significant effects for predicting the outcome of primary training are the FAA Exam Score and the Exam score on Stage 1 Exam. The Whole Model Test Report, shown in Figure 92, has a small p-value (<0.0001), which indicates that the overall model is significant, and an R^2 of 0.0515 indicating that 5.15 % of the variation in response is explained by the model.

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	7.89008	2	15.78017	0.0004*
Full	145.19930			
Reduced	153.08938			
	RSquare (U)		0.0515	
	AICc		296.44	
	BIC		309.477	
	Observations (or Sum Wgts)		578	

Figure 92. Logistics model whole model test for IFS

Figure 93 shows the confusion matrix for both the training and validation sets.

Training			Validation		
Actual	Predicted Count		Actual	Predicted Count	
Pri	NG	G	Pri	NG	G
NG	1	42	NG	0	16
G	0	535	G	1	245

Figure 93. Logistics model confusion matrix for IFS

As a result of the data imbalance, the model performs poorly for predicting Non-Graduates. To remedy this situation, and as discussed earlier, a cut off lower than 0.5 needs to be specified for classifying a predicted probability as Non-Graduates.

Figure 94 shows the prediction profiler. It indicates the predicted probability of the response variable PRI for each specified value of the predictor variables. As shown in the figure, a higher Exam score on Stage 1 Exam and IFS FAA Score, indicate a lower probability of Non-Graduate outcome in Advanced training.

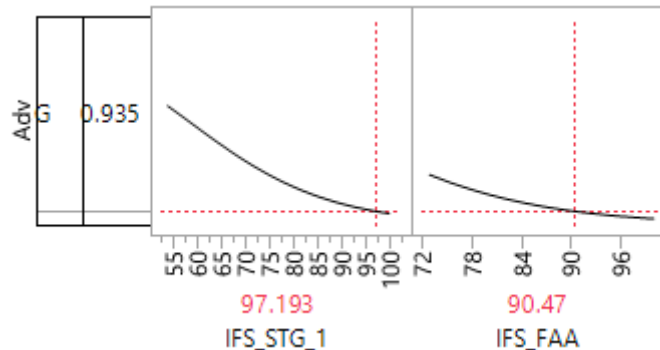


Figure 94. Logistics model prediction profiler for IFS

b. Bootstrap (Random) Forest Model

This model uses a bootstrap (random) forest to predict the outcome of advanced training from the full set of potential predictors (main effects) for those students who took the ASTBE. Figure 95 shows the overall model summary statistics. For the validation set, R^2 is -0.007 and the misclassification rate is 0.0706.

Measure	Training	Validation	Definition
Entropy RSquare	0.0924	-0.007	$1 - \text{Loglike}(\text{model}) / \text{Loglike}(0)$
Generalized RSquare	0.1152	-0.010	$(1 - (L(0)/L(\text{model}))^{2/n}) / (1 - L(0)^{2/n})$
Mean -Log p	0.2305	0.2571	$\sum -\text{Log}(p[j]) / n$
RMSE	0.2482	0.2575	$\sqrt{\sum (y[j] - p[j])^2 / n}$
Mean Abs Dev	0.1220	0.1282	$\sum y[j] - p[j] / n$
Misclassification Rate	0.0701	0.0706	$\sum (p[j] \neq p_{\text{Max}}) / n$
N	827	354	n

Figure 95. Bootstrap forest overall model summary for IFS

Figure 96 shows the column contribution of the different predictors with the FAA Exam Score (IFS_FAA), Exam score on Stage 1 Exam (IFS_STG_1), and the number of landings completed during IFS (IFS_TOTAL_LANDINGS) being the top three contributors.

Term	Number of Splits	G^2	Portion
IFS_FAA	5	6.62222083	0.2461
IFS_STG_1	5	5.32112722	0.1978
IFS_TOTAL_LANDINGS	6	3.96983972	0.1476
IFS_EOC	5	3.94141002	0.1465
IFS_FLT_FAIL	3	2.30081411	0.0855
IFS_STG_3	3	2.00988461	0.0747
IFS_TOTAL_FLIGHT_TIME	3	1.37367262	0.0511
IFS_STG_2	3	1.36434192	0.0507

Figure 96. Bootstrap forest model column contribution for IFS

Figure 97 shows the confusion matrix for both the training and validation sets.

Training			Validation		
Actual	Predicted Count		Actual	Predicted Count	
Pri	G	NG	Pri	G	NG
G	769	0	G	329	0
NG	58	0	NG	25	0

Figure 97. Bootstrap forest model confusion matrix for IFS

As indicated earlier and as a result of the data imbalance, the model does not predict any Non-Graduates although the misclassification rate is low. Similar to the approach used in the logistics regression model, we need to specify a probability cut off lower than 0.5 to classify more observations as Non-Graduates.

Figure 98 shows the prediction profiler. It indicates the predicted probability of the response variable PRI for each specified value of the predictor variables. As shown in the Figure 98, the response variable ADV has low sensitivity to changes in predictor variables.

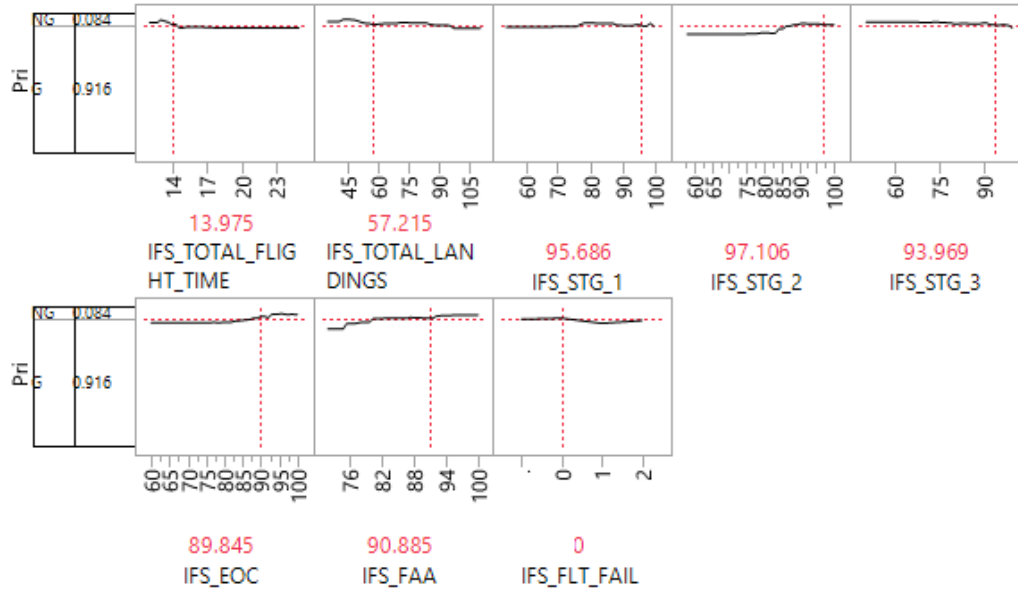


Figure 98. Bootstrap forest model prediction profiler for IFS

c. Model Comparison

We compare the developed models to select a best model. Figure 99 and Figure 100 compare these models for both the training and validation data sets. As shown in these figures, the Bootstrap Forest Model has a better R^2 for both the training and validation data sets. Note that because R^2 is defined as the proportion of variance explained by the fit, if the fit is actually worse than just fitting a horizontal line, as in the case of the validation data set, then R-square is negative.

Creator		Entropy RSquare	Generalized RSquare	Mean - Log p	RMSE	Mean Abs Dev	Misclassification Rate	N
Bootstrap Forest		0.0924	0.1152	0.2305	0.2482	0.1220	0.0701	827
Fit Nominal Logistic		0.0515	0.0655	0.2512	0.2571	0.1324	0.0727	578

Figure 99. Model comparison for training set for IFS

Creator		Entropy RSquare	Generalized RSquare	Mean - Log p	RMSE	Mean Abs Dev	Misclassification Rate	N
Bootstrap Forest		-0.007	-0.010	0.2571	0.2575	0.1282	0.0706	354
Fit Nominal Logistic		-0.026	-0.033	0.2359	0.2463	0.1268	0.0649	262

Figure 100. Model comparison for validation set for IFS

d. Conclusions

In this section, we developed a multiple logistics model and a bootstrap (random) forest model to predict the outcome of Advanced training (ADV) (Graduate or Non-Graduate/Attrite) from IFS data. Both models developed in this section do not seem to be a good fit as evident by the small R^2 values, which indicate that only a small percentage of the dependent variable variation is explained by the model. Based on the analysis of the developed models, we make the following conclusions:

- The two significant effects in the logistics model for predicting the outcome of primary training are the FAA Exam Score and the Exam score on Stage 1 Exam.
- For the bootstrap forest model, the top three variable contributors are the z-scores for the FAA Exam Score (IFS_FAA), Exam score on Stage 1 Exam (IFS_STG_1), and the number of landings completed during IFS (IFS_TOTAL_LANDINGS).
- The Bootstrap Forest Model has a better R^2 for both the training and validation data set.

3. Predicting Outcome of Advanced Training (ADV) from API Scores

In this section, we develop a multiple logistics model and a decision (partition) tree models to predict the outcome of advanced training (ADV) (Graduate or Non-Graduate/Attrite) from API data for students who took the ASTBE. API Predictors include the overall performance of API as Navy Standard Score and the number of academic failures in API.

a. Multiple Logistics Model

This multiple logistics model predicts the outcome of primary training from the set of potential predictors (main effects). Figure 101 shows the effect summary of the model.

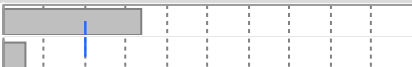
Source	LogWorth		PValue
API_NSS	3.372		0.00042
API_Test_FAILS	0.563		0.27379

Figure 101. Logistics model effect summary for API

As shown in Figure 101, the most significant effect for predicting the outcome of primary training is the overall performance of API as Navy Standard Score (API_NSS). The Whole Model Test Report, shown in Figure 102, has a small p-value (<0.0001), which indicates that the overall model is significant, and an R^2 of 0.0339 indicating that 3. % of the variation in response is explained by the model.

Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	7.12487	2	14.24974	0.0008*
Full	202.77335			
Reduced	209.89822			
	RSquare (U)		0.0339	
	AICc		411.576	
	BIC		425.693	
	Observations (or Sum Wgts)		825	

Figure 102. Logistics model whole model test for API

Figure 103 shows the confusion matrix for both the training and validation sets.

Training			Validation		
Actual	Predicted Count		Actual	Predicted Count	
Pri	NG	G	Pri	NG	G
NG	0	58	NG	0	25
G	0	767	G	0	329

Figure 103. Logistics model confusion matrix for API

As discussed earlier and as a result of the data imbalance, the model does not predict any Non-Graduates. To remedy this situation, we need to specify a cut off lower than 0.5 for classifying a predicted probability for Non-Graduates.

Figure 104 shows the prediction profiler. It indicates the predicted probability of the response variable PRI for each value of the API predictor variables. As shown in the figure, a higher value of the overall performance of API as Navy Standard score (API_NSS) and number of academic failures in API (API_Test_FAILS) is associated with a lower probability of Non-Graduate outcome.

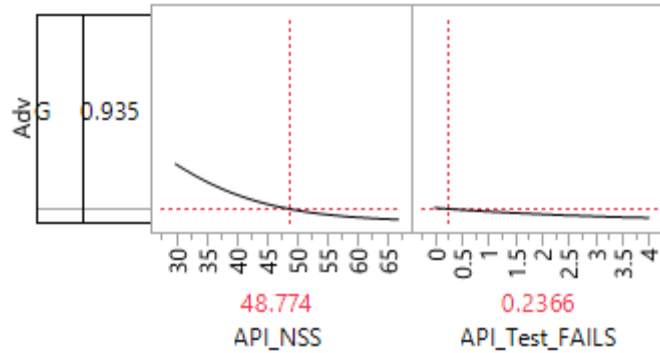


Figure 104. Logistics model prediction profiler for API

b. Decision (Partition) Tree Model

In this section, we fit a decision tree model to the API data for students who took the ASTBE Exam. Figure 105 shows the resulting decision tree pruned using the validation data to avoid overfitting. As shown in Figure 106, all splits occur on the API Navy Standard Score (API_NSS), which is the most significant predictor. The fit details and confusion matrices are shown in Figure 107, and the leaf report, shown in Figure 108, displays the rules generated by the decision tree.

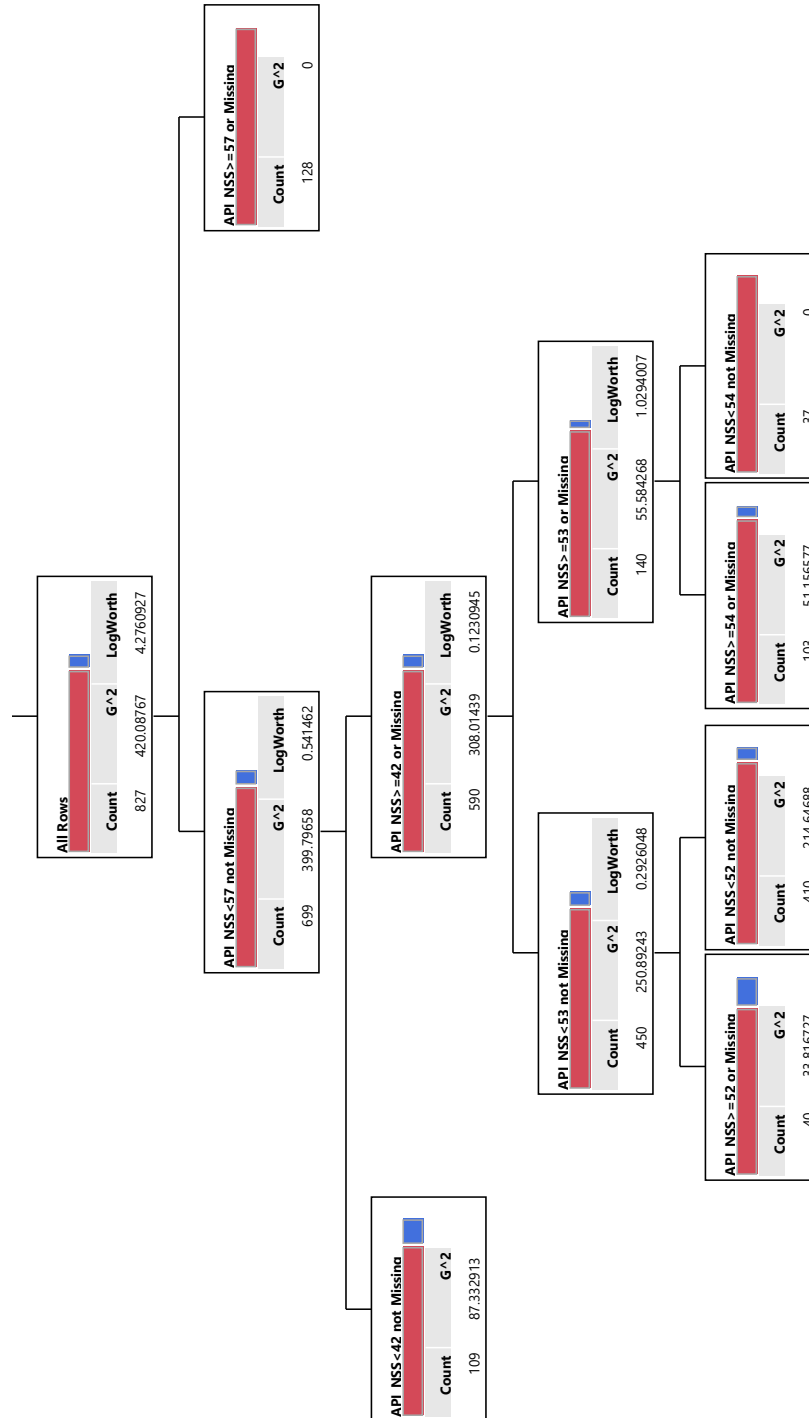


Figure 105. Decision tree for API

Term	Number of Splits	G ²	Portion
API_NSS	5	33.1345723	1.0000
API_Test_FAILS	0	0	0.0000

Figure 106. Decision tree model column contribution for API

Measure	Training	Validation	Definition
Entropy RSquare	0.0782	0.0908	$1 - \text{Loglike}(\text{model}) / \text{Loglike}(0)$
Generalized RSquare	0.0978	0.1133	$(1 - (L(0)/L(\text{model}))^{2/n}) / (1 - L(0)^{2/n})$
Mean -Log p	0.2341	0.2321	$\sum -\text{Log}(p[j]) / n$
RMSE	0.2516	0.2507	$\sqrt{\sum (y[j] - \rho[j])^2 / n}$
Mean Abs Dev	0.1267	0.1303	$\sum y[j] - \rho[j] / n$
Misclassification Rate	0.0701	0.0706	$\sum (\rho[j] \neq p\text{Max}) / n$
N	827	354	n

Training			Validation		
Actual	Predicted Count		Actual	Predicted Count	
Pri	G	NG	Pri	G	NG
G	769	0	G	329	0
NG	58	0	NG	25	0

Figure 107. Decision tree model fit details and confusion matrices for API

Leaf Label	G	NG
API_NSS<57 not Missing&API_NSS>=42 or Missing&API_NSS<53 not Missing&API_NSS>=52 or Missing	0.8519	0.1481
API_NSS<57 not Missing&API_NSS<42 not Missing	0.8630	0.1370
API_NSS<57 not Missing&API_NSS>=42 or Missing&API_NSS<53 not Missing&API_NSS<52 not Missing	0.9268	0.0732
API_NSS<57 not Missing&API_NSS>=42 or Missing&API_NSS>=53 or Missing&API_NSS>=54 or Missing	0.9320	0.0680
API_NSS<57 not Missing&API_NSS>=42 or Missing&API_NSS>=53 or Missing&API_NSS<54 not Missing	0.9982	0.0018
API_NSS>=57 or Missing	0.9995	0.0005

Figure 108. Decision tree model leaf report for API

c. Model Comparison

We compare the two developed models to select a best model. Figure 109 and Figure 110 compare these models for both the training and validation data sets. As shown in these figures, the Decision (Partition) Model has a better R^2 for both the training and the validation data sets.

Creator		Entropy RSquare	Generalized RSquare	Mean - Log p	RMSE	Mean Abs Dev	Misclassification Rate	N
Partition		0.0782	0.0978	0.2341	0.2516	0.1267	0.0701	827
Fit Nominal Logistic		0.0339	0.0429	0.2458	0.2536	0.1285	0.0703	825

Figure 109. Model comparison for training set for API

Creator		Entropy RSquare	Generalized RSquare	Mean - Log p	RMSE	Mean Abs Dev	Misclassification Rate	N
Partition		0.0908	0.1133	0.2321	0.2507	0.1303	0.0706	354
Fit Nominal Logistic		0.0775	0.0970	0.2355	0.2510	0.1329	0.0706	354

Figure 110. Model comparison for validation set for API

d. Conclusions

In this section, we developed a multiple logistics model and a decision tree model to predict the outcome of advanced training (ADV) (Graduate or Non-Graduate/Attrite) from API data. Both models developed in this section do not seem to be a good fit as evident by the small R^2 values of the validation set (<10%), which indicate that only a small percentage of the dependent variable variation is explained by the model. Based on the analysis of the developed models, we make the following conclusions:

- The most significant effect in the logistics model for predicting the outcome of primary training is the overall performance of API as Navy Standard Score (API_NSS).
- Similarly, most of the splits for the decision tree model occur on the API Navy Standard Score (API_NSS), indicating it is the most significant predictor.
- The decision tree (partition) model is the better model (higher R^2) for both the training and validation data sets.

THIS PAGE INTENTIONALLY LEFT BLANK

VI. SUMMARY AND CONCLUSIONS

A. SUMMARY

In this research effort we developed several models for predicting the graduation status (Graduate or Non-Graduate/Attrite) of student aviators in primary and advanced flight training from ASTB, IFS, and API data for both ASTB345 and ASTBE students. The developed models included multiple logistics, decision tree, bootstrap (random) forest, and generalized (penalized) linear models. It is important to note that the R^2 of all developed models is small (<10%) indicating that only a small amount of the variation in the outcome is explained by the model. The following is a summary of the results of the developed models:

1. Predicting outcome of PRI training using ASTB345 Data

a. From ASTB345 Scores

- The top three significant effects in the logistics model are the z-scores for the Academic Qualifications Rating (AQR_RAW_345), Officer Aptitude Rating (OAR_RAW_345), and Aviation and Nautical Information (ANI_RAW_345).
- The top three significant effects in the random forest model are the z-scores for the Flight Officer Aptitude Rating (FOFAR_RAW_345), Spatial Awareness Rating (SAT_RAW_345), and Academic Qualifications Rating (AQR_RAW_345).
- The top significant effects for the Lasso model, Elastic Net, and Ridge is the z-score for the Flight Officer Aptitude Rating (FOFAR_RAW_345).
- Demographic variables don't seem to be significant in predicting the outcome variable.
- Due to collinearity among predictor variables, tree and penalized models would be more accurate models. In fact, model comparison on validation data indicated the three generalized linear models performed the best.

b. From IFS Scores

- The top three significant effects in the logistics model are the number of flight failures in IFS, the FAA exam score, and exam score on Stage 1.

- For bootstrap forest model, the top three variable contributors are number of hours of flight time received during IFS (IFS_TOTAL_FLIGHT_TIME), FAA Exam Score (IFS_FAA), and number of landings completed during IFS (IFS_TOTAL_LANDINGS).
- The most significant predictors in the Elastic Net model are the number of flight failures in IFS (IFS_FLT_FAIL), Exam Score on Stage 1 Exam (IFS_STG_1), and FAA Exam Score (IFS_FAA).
- The Elastic Net model has the best R^2 on validation data followed by the bootstrap forest model and the logistics model.

c. From API Scores

- The most significant effect in the logistics model is the overall performance of API as Navy Standard Score (API_NSS).
- Similarly, for the decision tree model, the most significant variable for predicting the outcome of primary training is the Navy Standard Score (API_NSS).
- The logistics model and decision tree model have comparable R^2 on both the training and validation data sets.

2. Predicting outcome of PRI training using ASTBE Data

a. From ASTBE Scores

- For the multiple logistics model, the top three significant effects are the z-scores for Officer Aptitude Rating on ASTBE (OAR_Z_ASTBE), Personality 9 Rating (Personality_9_Z_ASTBE), and Dichotic Listening Task on ASTBE (DLTFactor_Z_ASTBE).
- For bootstrap forest model, the top three variable contributors are the z-scores for Personality 9 Rating (Personality_9_Z_ASTBE), Dichotic Listening Task (DLTFactor_Z_ASTBE), and Academic Qualifications Rating (AQR_Z_ASTBE).

b. From IFS Scores

- The top three significant effects in the logistics model are the number of flight failures in IFS, the End of Course Exam Score, and the FAA Exam Score.

- For bootstrap forest model, the top three variable contributors are number of landings completed during IFS (IFS_TOTAL_LANDINGS), End of Course Exam Score (IFS_EOC), and FAA Exam Score (IFS_FAA).
- The bootstrap forest model performed better on both the training and validation data.

c. API Scores

- For the logistics model, the most significant effect is the overall performance of API as Navy Standard Score (API_NSS).
- Similarly, most of the splits for the decision tree model occur on the API Navy Standard Score (API_NSS), indicating it is the most significant predictor.

3. Predicting outcome of ADV training using ASTBE Data

a. From ASTBE Scores

- For the multiple logistics model, the top two significant effects for predicting the outcome of advanced training are the z-scores for Academic Qualifications Rating on ASTBE (AQR_Z_ASTBE) and Officer Aptitude Rating on ASTBE (OAR_Z_ASTBE).
- For bootstrap forest model, the top three variable contributors are the z-scores for the Personality 9 Rating (Personality_9_Z_ASTBE), Academic Qualifications Rating on ASTBE (AQR_Z_ASTBE), and Dichotic Listening Task (DLTFactor_Z_ASTBE).
- The Bootstrap Forest Model has a better R^2 for the training data set, while the Nominal Logistics Model has a better R^2 for the validation data set.

b. From IFS Scores

- The two significant effects in the logistics model for predicting the outcome of advanced training are the FAA Exam Score and the Exam score on Stage 1 Exam.
- For bootstrap forest model, the top three variable contributors are the z-scores for the FAA Exam Score (IFS_FAA), Exam score on Stage 1 Exam

(IFS_STG_1), and the number of landings completed during IFS (IFS_TOTAL_LANDINGS)

- The Bootstrap Forest Model has a better R^2 for both the training and validation data set

c. From API Scores

- The most significant effect in the logistics model is the overall performance of API as Navy Standard Score (API_NSS).
- Similarly, most of the splits for the decision tree model occur on the API Navy Standard Score (API_NSS), indicating it is the most significant predictor.
- The decision tree (partition) model is the better model (higher R^2) for both the training and validation data sets.

B. CONCLUSIONS

Based on the results of the developed models, we make the following general conclusions: 1) ASTB (ASTB345 and ASTBE) scores have very little or no effects on predicting graduation or non-graduation in primary and advanced training, 2) ASTB345 scores are strongly correlated, while the ASTBE scores have generally a weak to medium correlation, 3) For ASTB345 scores, the most important predictor variable seems to be the z-score for the Flight Officer Flight Aptitude Rating (FOFAR_RAW_345), 4) Most significant predictors for ASTBE scores include the z-scores for Academic Qualifications Rating on ASTBE (AQR_Z_ASTBE), Officer Aptitude Rating on ASTBE (OAR_Z_ASTBE), Personality 9 Rating (Personality_9_Z_ASTBE), and Dichotic Listening Task on ASTBE (DLTFactor_Z_ASTBE), 5) Demographic variables don't seem to be significant in predicting the outcome variable, 6) The top two significant predictors for graduation or non-graduation in primary and advanced training in IFS data are the number of flight failures in IFS and the FAA exam score, and 7) The most significant predictor in the API data for predicting the outcome of primary and advanced training is the Navy Standard Score (API_NSS).

LIST OF REFERENCES

- Abbott, D. (2014). *Applied Predictive Analytics: Principles and Techniques for the Professional Data Analyst*. Indianapolis: John Wiley & Sons, Inc.
- Chief of Naval Air Training (n.d.). *Naval Aviator Training*. Retrieved August 3, 2020, from <https://www.cnatra.navy.mil/training.asp>
- EMC Education Services. (2015). *Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data*. Indianapolis: John Wiley & Sons, Inc.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.). Springer-Verlag.
- Klimbert, R., & McCullough, B. D. (2016). *Fundamentals of Predictive Analytics with JMP®*, Second Edition (2nd ed.). SAS Institute.
- Natali, M. (2019). *Naval aviation order of training* [Unpublished manuscript].
- Navy Medicine Operational Training Center (NMOTC) – Pensacola (n.d.). *ASTB-E Overview*. Retrieved August 10, 2020, from <https://www.med.navy.mil/sites/nmotc/nami/pages/astboverview.aspx>
- SAS Institute Inc. (2018). *JMP® 14 Fitting Linear Models*. Cary, NC: SAS Institute Inc.
- SAS Institute Inc. (2018). *JMP® 14 Predictive and Specialized Modeling*. Cary, NC: SAS Institute Inc.
- Vorhies, W. (2016, July 26). *CRISP-DM – a Standard Methodology to Ensure a Good Outcome*. Data Science Central. <https://www.datasciencecentral.com/profiles/blogs/crisp-dm-a-standard-methodology-to-ensure-a-good-outcome>

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
3. Research Sponsored Programs Office, Code 41
Naval Postgraduate School
Monterey, CA 93943
4. CAPT Scott Starkey
Naval Air Training Command (NATRACOM)
Naval Air Station Corpus Christi
Corpus Christi, TX 78419
5. LT Michael Natali
Naval Air Training Command (NATRACOM)
Naval Air Station Corpus Christi
Corpus Christi, TX 78419