INSTITUTE FOR DEFENSE ANALYSES

Retention Prediction Model - Army

WEAI 2023
Defense Sessions

Alan Gelder
Jordan Marcusse
Ed Wang
Erin Eifert
Vivian Gao

July 2023
Approved for public release;
distribution is unlimited.
IDA Document NS D-33558
Log: H 23-000227

INSTITUTE FOR DEFENSE ANALYSES
730 E.Glebe Rd
Alexandria, VA 22305

The Institute for Defense Analyses is a nonprofit corporation that operates three Federally Funded Research and Development Centers. Its mission is to answer the most challenging U.S. security and science policy questions with objective analysis, leveraging extraordinary scientific, technical, and analytic expertise.

About this Publication
The work was conducted by the Institute for Defense Analyses (IDA) under CRP C6627. The views, opinions, and findings should not be construed as representing the official position of either the Department of Defense or the sponsoring organization.

Acknowledgments
The authors thank Joe Coombs for expert editing; and Amberlee Mabe-Stanberry for proficiently producing this document.

For More Information:
Dr. Jordan R. Marcusse, Leader Author
jmarcuss@ida.org, 703-845-2147
Ms. Jessica L. Stewart, Director, SFRD
jstewart@ida.org, 703-575-4530

INSTITUTE FOR DEFENSE ANALYSES

IDA Document NS D-33558

Retention Prediction Model - Army

WEAI 2023
Defense Sessions

Alan Gelder
Jordan Marcusse
Ed Wang
Erin Eifert
Vivian Gao

This page is intentionally blank.

Executive Summary

This presentation introduces the Retention Prediction Model – Army (RPM-A), a machine learning tool that forecasts individual soldier retention. The Institute for Defense Analyses (IDA) originally developed a version of the Retention Prediction Model in 2018 as part of a project for the Office of the Under Secretary of Defense for Personnel and Readiness, Military Personnel Policy. IDA is now delivering this model within the Army's Person-Event Data Environment (PDE) along with code to facilitate updates to the data, model, and forecasts, as well as a dashboard to support the model's use by Army personnel.

Among the sources of information that inform the model are demographics, family, career and pay, unit characteristics, casualties, deployments, the external job market, and performance data.

In this presentation, we discuss the methodology, the data inputs, and the tools we produced as part of this effort.

This page is intentionally blank.

Retention Prediction Model - Army

WEAI 2023

Defense Sessions

Alan Gelder

Erin Eifert

Jordan Marcusse

Vivian Gao

Ed Wang

3 July 2023

The Retention Prediction Model - Army (RPM-A)

Goal: Forecast future retention and build tools for Army leaders to leverage the forecasts

Approach: Use machine learning toolkit to forecast retention for individual service members

- Stand up IDA's in-house personnel data pipeline and expand it to current and historical Army personnel data within the cloud version of Army's Person-Event Data Environment (PDE)
- Train a machine learning model on these data to produce retention forecasts for every member of the Army
- Host output forecasts in a dashboard for Army leaders

Example use case: Identify likely future shortfalls within populations such as occupation, rank, and performance.

Status: IDA has stood up the RPM-A for officers in PDE, including forecasts and a dashboard, as well as a model for the full force trained on a more limited set of data

Planned topics for today's discussion

The RPM-A and underlying analytical methods

Hosting RPM-A within Army's environment

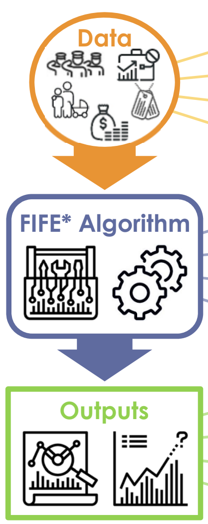
Analysis of data inputs

Model performance

IDA

The RPM delivers person-level retention forecasts

OUSD(P&R) seeded the development of this flexible toolkit for retention analysis



- Personnel records
- Family information
- Economic conditions
- Other pertinent data

- Adaptable, modular
- Open-source
- Machine learning
- Survival analysis

- Person-level forecasts
- Group forecasts
- Dashboard inputs
- Launching in PDE

Research applications for Services and OSD

- **Army:** Full modeling pipeline in Army systems
- **ANG:** Forecast and program for training slots
- **Navy:** Identify correlates of officer exit & promotion
- **P&R:** Examine Academy graduate ADSOs**
- **USU:** Measure return on USU physician training
- **QRMC:** Estimate effects of past compensation

Partnerships expand FIFE & RPM capabilities

- Operationalize in shared/service environments - PDE
- Link forecasts to characteristics of interest
- Enhance and expand data development

IDA

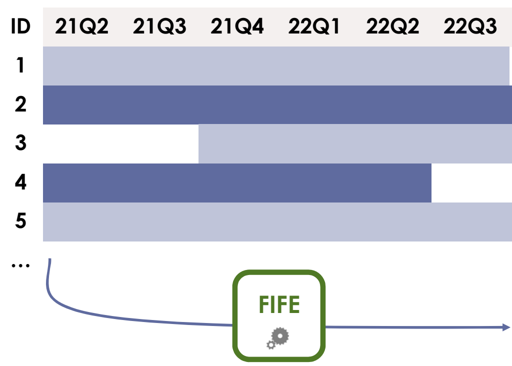
*Finite Interval Forecasting Engine **ADSO: Active Duty Service Obligation

FIFE estimates retention with machine learning, fitting a binary prediction model to each future time horizon

Input: Imbalanced panel data on individuals

Training data: X = Individual data at each time period

Y = Retention 1 period ahead, Retention 2 periods ahead, …



	P(observation)		
ID	23Q3	23Q4	...
2	0.83	0.79	...
3	0.98	0.97	...
...	...	...	...

IDA

FIFE's methodology is more flexible than prevailing techniques

Traditional tools for survival analysis

Kaplan-Meier: $H(t)$—share surviving to time horizon t

Proportional Hazards: $H(t)\,f(x)$—now a function of feature values

Our method: $H(t, x)$—allows interactions with time and features

We effectively compute $f_t(x)$ for each forecast horizon, where

$$f_t(x) = P(\text{Remain from } t\text{-}1 \text{ to } t \mid \text{Remained in sample from } 0 \text{ to } t\text{-}1, x)$$

$$H(t, x) = f_1(x)\, f_2(x) \ldots f_t(x)$$

IDA

Classification method - choice of $f_t(x)$

RPM-A uses **gradient-boosted trees** (via LightGBM)

Benefits	Drawbacks
Best performance among classifiers Relatively stable performance across hyperparameter specifications and data inputs Fast computation	Cannot perform mathematical operations across features (x - y > a), time trends, … Feature engineering helps to resolve some of these issues

Alternative classifiers available in FIFE

Feed-forward neural network (via Keras)
Proportional hazards (via a constrained neural network)
Group rates (via pandas *groupby*; tantamount to fixed effects)

IDA

Hosting RPM-A within Army's environment

IDA

The RPM-A in Army's Person-Event Data Environment

Data inputs to RPM-A

- Demographics
- Dependents
- Career and pay
- Unit traits
- Casualty
- Deployments
- External job market
- Loss categories
- **Performance**
- **Rater characteristics**
- **Fitness**
- **Drug testing**

Bolded data inputs are Army-specific

Operationalizing the RPM-A puts the full pipeline in Army's hands

Data preparation, modeling, and dashboard reside in the Person-Event Data Environment (PDE)

Army can use the retention forecasts within a business intelligence system

Army can directly control access to and applications of RPM-A outputs

IDA

We streamlined the data-modeling-analysis pipeline

1. Query raw data » 2. Clean data » 3. Engineer features » 4. Merge datasets

1. Query raw data
- Access data from Army, DoD, and other sources

2. Clean data
- Standardize missing values
- Map recoded values
- Merge across reformats
- Other cleaning processes

4. Merge datasets
- Merge tables
- Store in quickly read, column-based format (.feather)

3. Engineer features

Examples

Dates → Time since/until event
- Officer has 748 days until her ADSO date

Past events → Event histories
- Officer's most recent senior rater evaluation was "Most Qualified"
- Officer was deployed for 271 days in the past 12 months

Descriptions of dependents → Family structure
- Officer has 2 children under the age of 6

Status → Change in status
- Officer got divorced in the past year

Aggregate data from other units of observation
- Unit/Rater/Military Occupational Specialties (MOS) characteristics

IDA

Model training

With this efficient data pipeline, we can train versions of the model during scheduled high-compute days

We assess model performance and feature importance across a variety of specifications, such as:

- Feature inputs
- Training period
- Forecast date

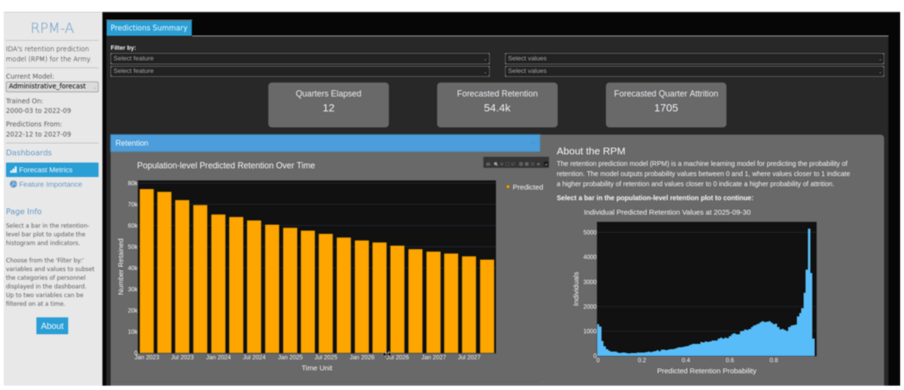
To train a new model: specify a population, feature inputs, time period, and hyperparameters

We train a set of predefined models and store their forecasts as quarterly data updates arrive

- Administrative RPM-A
- Research RPM-A (includes data with restrictions on use or limited dates)

IDA

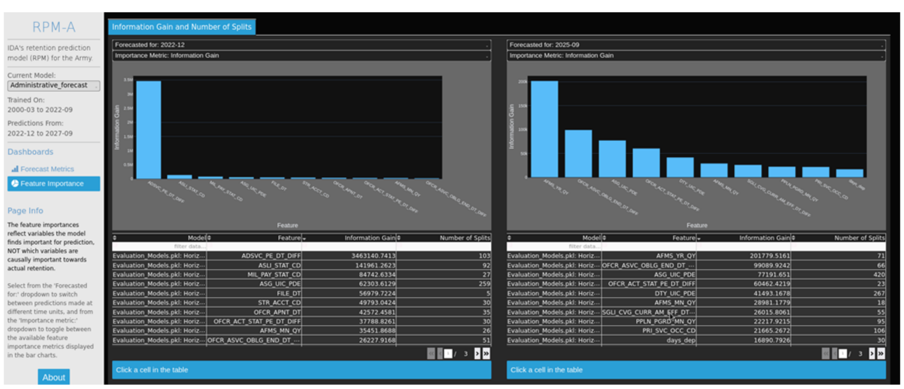
The RPM-A dashboard supports analysis and planning



IDA

Dashboard uses: Aggregate forecasts over user-defined populations

Dashboard uses: Identify data elements coincident with retention/exit



Dashboard uses: Evaluate performance of forecasts from prior periods against actual outcomes

IDA

Considerations for applications of the RPM-A

The RPM-A forecasts assume that patterns in the past will continue

The RPM-A cannot identify the causal impact of data elements

Using forecasts for individual-level career management is risky

- Accuracy may be insufficient for person-level actions
- Even with perfect accuracy, there are concerns of ethics, fairness, creation of unintended incentives, and other matters

The RPM-A's retention probabilities can supplement other analyses

IDA

Model performance

IDA

16

Model Performance: All officers, 3 years

Predictions of retention from September 2019 to September 2022:

DMDC and economic data, or **all data**

With a prediction threshold of 50% for assignment to remain/stay[1]

Retention rate: 75.4%

		Outcome: Exit	Outcome: Stay
Forecast	Exit	7,217 7,495	2,300 2,125
Forecast	Stay	12,111 11,833	56,797 56,972

Statistics[2]	
Accuracy $\frac{\{\text{forecast = outcome}\}}{\{\text{all observations}\}}$	81.6% 82.2%
Precision $\frac{\{\text{outcome = forecast = exit}\}}{\{\text{forecast = exit}\}}$	75.8% 77.9%
Recall $\frac{\{\text{outcome = forecast = exit}\}}{\{\text{outcome = exit}\}}$	37.3% 38.8%
AUROC $\int_{\text{threshold}=0}^{1} \text{accuracy}(\text{threshold})$	0.812 0.824

[1] Other thresholds trade off false positives and false negatives, improving precision or recall, and may be more appropriate for specific use cases

[2] Accuracy: rate of actual exits & stays being identified; precision: rate of forecasted exits being correct; recall: rate of actual exits being identified

AUROC: Area Under the Receiver Operating Characteristic

IDA

Model Performance: 1 year

17

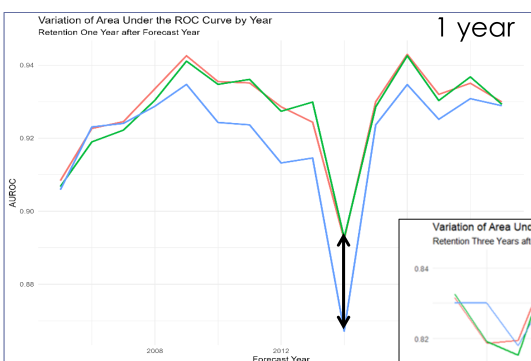
Select informative features by information gain

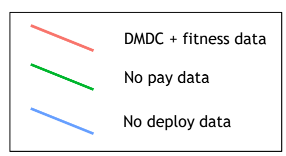
1-4 quarters from forecast date	9-12 quarters from forecast date
Days until projected end of service	Years of active federal military service
Assigned UIC (Unit Identification Code)	Assigned UIC
Active Service Loss Incentive status	Days deployed in career
Duty UIC	Duty UIC
Strength accounting code	Days until projected end of service
Years of active federal military service	Pay grade
Days deployed in career	Occupation code (MOS)
Date	ZIP code of home address
Pay status	Quarter
Days spent in pay grade	Days spent in pay grade

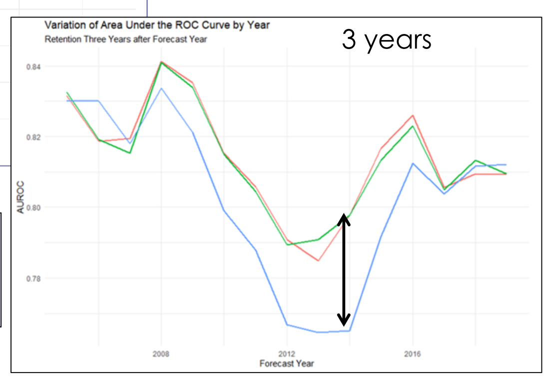
Army-specific data elements in top 10-100 features:
- Most recent Officer Evaluation Report (OER) was referred
- Ever had a referred OER
- Reason not rated on most recent OER
- Box check on most recent OER
- Percent of box checks "Most qualified" on OER
- Ever had a profile on OER
- Fitness score percentile
- 2 mile run percentile

IDA

Model performance, input importance vary over time







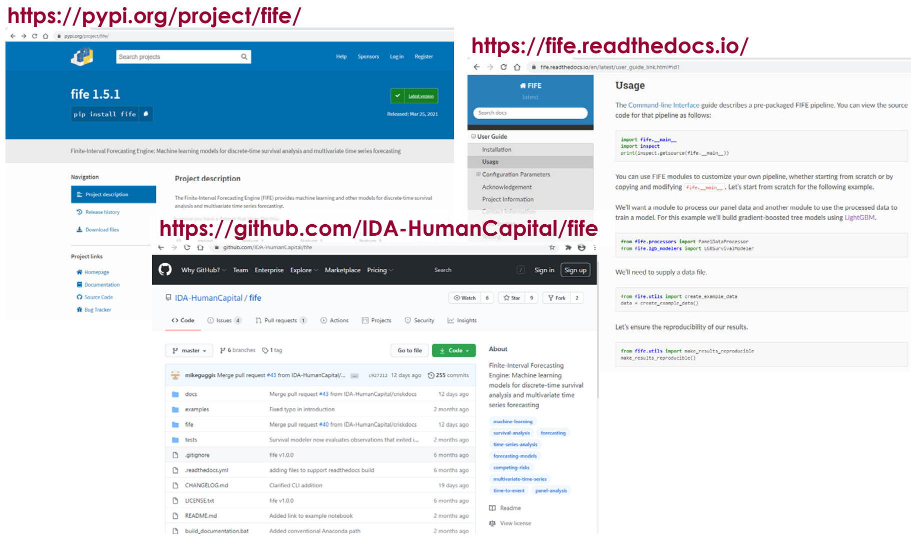
Deployment activity was an important predictor from 2008-2016, but is much less of a predictor now

Model performance of RPM-A trained on DMDC (Defense Manpower Data Center) + fitness data for all officers, trained on rolling 4-year windows of data preceding the forecast date

IDA

Backup

FIFE is available in the open source



What’s in the FIFE package?

Panel Data Processor

Computes survival durations, identifies censorship, drops degenerate and duplicate features, and identifies training/validation sets

Survival Modelers

Can select from gradient-boosted trees (via LightGBM), feed-forward neural network (via Keras), proportional hazards, or group rates

State Modelers

Computes the future value of a feature conditional on survival

Exit Modelers

Computes competing risk of exit under various conditions

Feature Importance Attribution

Identifies the change in predictive power using SHAP analysis

IDA

Model Performance: All officers, 1 year

Predictions of retention from September 2019 to September 2020:

DMDC and economic data, or **all data**

With a prediction threshold of 50% for assignment to remain/stay[1]

Retention rate: 92.5%

		Outcome	
		Exit	Stay
Forecast	Exit	3,455 3,533	336 400
	Stay	2,545 2,467	72,059 72,025

Statistics[2]	
Accuracy $\frac{\{\text{forecast = outcome}\}}{\{\text{all observations}\}}$	96.2% 96.3%
Precision $\frac{\{\text{outcome = forecast = exit}\}}{\{\text{forecast = exit}\}}$	90.4% 89.8%
Recall $\frac{\{\text{outcome = forecast = exit}\}}{\{\text{outcome = exit}\}}$	57.6% 58.9%
AUROC $\int_{\text{threshold}=0}^{1} \text{accuracy}(\text{threshold})$	0.925 0.933

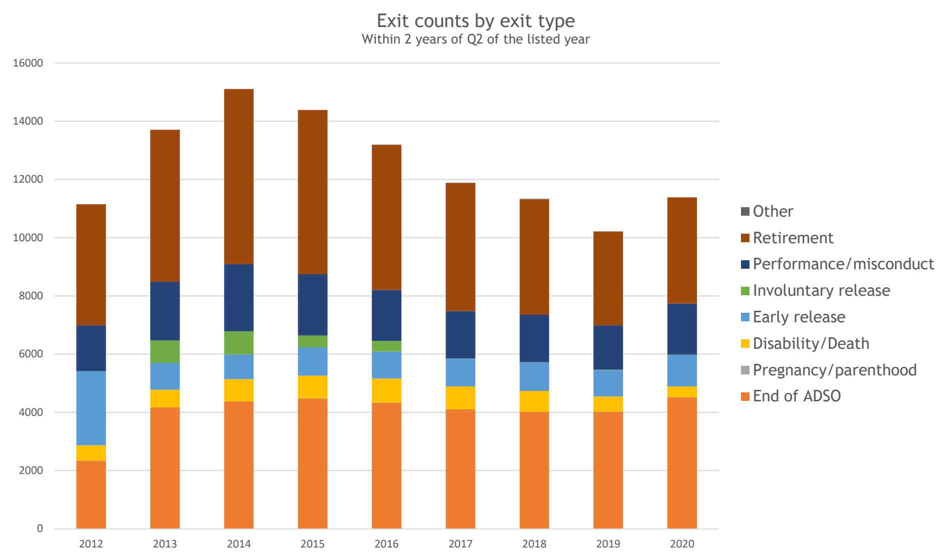
AUROC: Area Under the Receiver Operating Characteristic

[1] Other thresholds trade off false positives and false negatives, improving precision or recall, and may be more appropriate for specific use cases
[2] Accuracy: rate of actual exits & stays being identified; precision: rate of forecasted exits being correct; recall: rate of actual exits being identified

IDA

Model Performance: 3 years

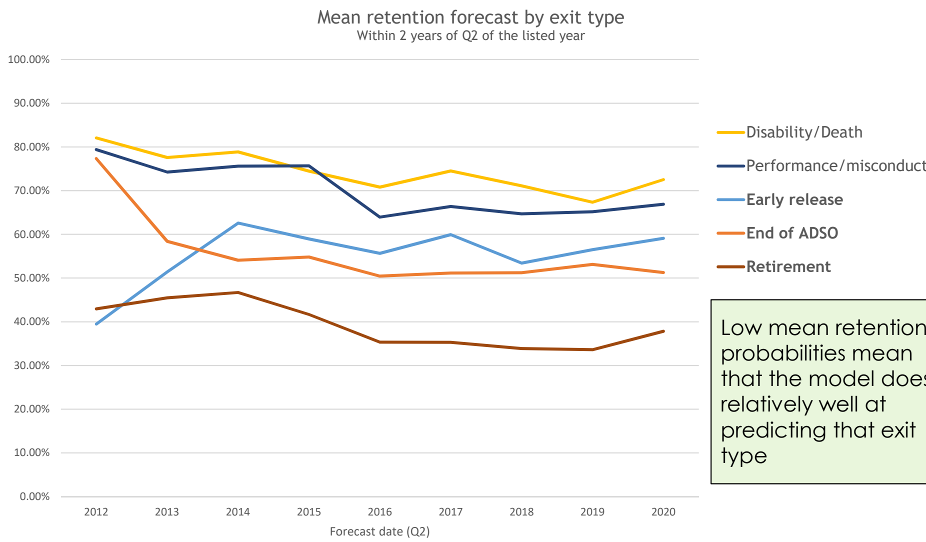
Frequency of exit types



IDA

Mean retention forecast by types of subsequent exit

2 year-ahead retention forecasts for officers who exited within 2 years of the forecast date

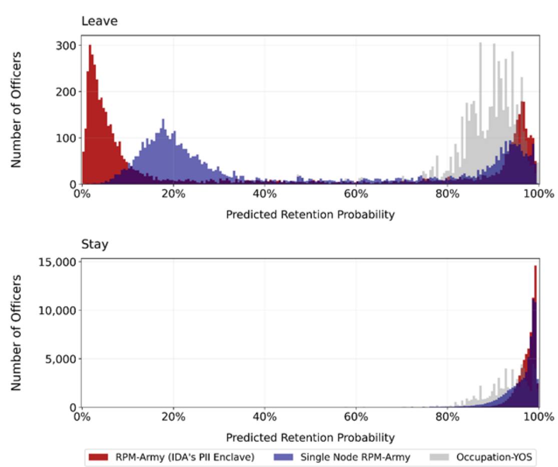


Low mean retention probabilities mean that the model does relatively well at predicting that exit type

IDA

A core set of features capture a large amount of the predictive power over a 1-year time horizon

Population:	All officers
Forecast date:	December 2019
Retention to:	December 2020
Training data:	March 2000- September 2019



	Accuracy	AUROC
DMDC and econ data	96.6%	0.923
28 selected fields	94.8%	0.918

Exit/Remain Classification Threshold = 0.5

IDA

Complete list of features included in Single Node RPM-A

28 Data Input Fields

Direct Indicators of Exit
- Active Duty Service Projected End Date
- Strength Accounting Code
- Personnel Strength Status Code
- Active Service Loss Incentive

Unit Information
- TOE/TDA type of Assigned Unit
- TOE/TDA type of Assigned Unit
- Assigned Unit Major Command Group
- Assigned Base
- Duty Base

Military Career
- Primary AOC
- Secondary AOC
- Duty AOC
- Years in Paygrade
- Months of Military Experience
- Current/previous command
- Component

Education/Quality
- Education Level
- Joint Professional Military Education
- Professional Military Education
- Source of Accession

Demographics
- Age
- Gender
- Race
- Ethnicity
- Source of U.S. citizenship

Return

IDA

This page is intentionally blank.