



Using Machine Learning Effectively

Machine learning (ML) is rapidly becoming intricately woven into many facets of government and industry decision-making, offering leaders broad opportunities to use ML techniques and other “big data” tools. When applied and interpreted appropriately, the results can inform and improve organizational decisions.

ML techniques and other big data tools offer more flexibility than traditional statistical modeling by identifying and making use of complex nuanced patterns and interactions within data. To enable successful use of ML, IDA recommends that leaders pursue the following requirements:

- **Facilitate data access** by reducing procedural and technical barriers to data access and sharing across both inter- and intra-organizational boundaries.
- **Improve data quality** by implementing a pipeline for transforming “raw data” objects into “research ready” data objects that are cleaned, standardized, documented, and reproducible.
- **Enhance data scope** by archiving data over long periods (to capture processes taking years or decades) and documenting both past and present policies and how they affect decisions.
- **Invest in open, reusable tools** by developing and sustaining reusable toolkits of algorithms and data and encouraging an appropriate open-source, collaborative culture for analytic tools.
- **Audit for legal, moral, and ethical use** by adopting best practices for auditing the results and use of operational ML tools (for example, to avoid perpetuating social or other biases that may be present in the data).

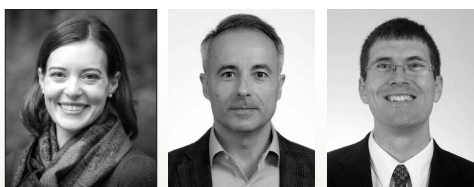
Careful matching of analytic questions to techniques is critical to obtaining meaningful results. Questions that can be addressed with ML generally fall into three main categories. The table below illustrates these categories, with examples from defense personnel management.

	What?	How or Why?	What if?
Objective	Describe current or past conditions and forecast future outcomes based on <i>existing trends</i>	Assess how a previous <i>event</i> or <i>enacted policy</i> affects a given outcome	Assess how a future <i>event</i> or <i>potential policy</i> affects a given outcome, if it were to occur
Benefits of ML adoption	Enhances the accuracy and sophistication of patterns and predictions that are identified from the data	Provides an understanding of complex impact mechanisms, identifies prominently affected populations, and produces individual-specific effects	Improves the scalability, speed, and complexity of traditional algorithms. Provides a richer understanding of impact mechanisms and individual-specific effects
Example questions	<p>What career fields are likely to have staffing shortfalls in 5 years?</p> <p>What units are likely to maintain strong cohesion and morale?</p> <p>Which candidates are likely to succeed in a particular training program?</p> <p>Which service members are likely to renew their contracts?</p>	<p>How did the Navy's Assignment Incentive Pay (AIP) affect retention of individuals in hard-to-fill billets?</p> <p>How did the COVID-19 pandemic affect mission-capable rates?</p> <p>Why are reports about sexual misconduct increasing?</p>	<p>If Selective Retention Bonuses (SRBs) were set at a given level, what would be the effect on reenlistment rates?</p> <p>If Navy sea tours were extended, what would be the effect on performance and health?</p> <p>If Congress enacted a new retirement benefit, what would be the effect on retention?</p>

Producing high-quality strategic and operational decisions informed by ML techniques demands careful research question framing, rich and reliable data, proper application of appropriate analytic tools, and accurate interpretation of results. To accomplish this—and avoid mistaken or misleading results—government and industry decision-makers must:

- Understand what the various ML techniques can and cannot provide;
- Choose techniques that are appropriate for answering their specific questions; and
- Interpret the results within the context of the methods and data.

Improved analytics and operational outcomes are possible when leaders are informed on how to interpret ML results, and when leaders establish an analytic culture with appropriate skills and knowledge, robust methodological and data quality control, and sufficient computational and data resources. With these requirements in place, leaders can employ ML capabilities to their full potential.



This publication was produced by a team from the Human Capital Group (HCG) in the Strategy, Forces and Resources Division of IDA's Systems and Analyses Center. The team included HCG leader **Julie A. Lockwood** (formerly Pechacek), jlockwood@ida.org, **Dave I. Cotting**, dcotting@ida.org, and **Alan B. Gelder**, agelder@ida.org.

Based on **IDA P-13174**, *Leveraging Machine Learning in Defense Personnel Analyses*, J. A. Lockwood, A. B. Gelder, M. S. Goldberg, J. T. Brooks, and G. M. Prugh, May 2020. Research was sponsored by the Office of the Under Secretary of Defense for Personnel and Readiness.