SYM-AM-17-044

Proceedings
of the
Fourteenth Annual Acquisition Research Symposium

Wednesday Sessions
Volume I

Acquisition Research:
Creating Synergy for Informed Change

April 26–27, 2017

Published March 31, 2017

Prepared for the Naval Postgraduate School, Monterey, CA 93943.

Data Consolidation of Disparate Procurement Data Sources for Correlated Performance-Based Acquisition Decision Support

Samantha Nangia—is a PMP and DAWIA-certified Project Manager and Business Technology Subject Matter Expert (SME). In her current role, she helps manage and oversee the suite of Acquisition Technology Solutions for the Department of the Navy (DoN) on the staff of the ASN(RD&A) DASN (AP) eBusiness Policy and Oversight (eBPO) Division. She has been dubbed the Navy's Procurement Data SME and is working to transform the current set of legacy Navy contract reporting and BI capabilities into a single data-rich enterprise solution. Nangia received her bachelor's degree in Computer Science from Rollins College in Winter Park, FL, in 2002 and her master's degree in Business Administration from Georgetown University in 2009.

Ryan Dickover—PhD, currently serves on the ASN(RD&A) staff as Director of DASN (AP)'s eBPO division. This division is responsible for the integration and governing of the suite of information systems used to acquire over $150 billion in DoN assets annually. Dr. Dickover is a member of the Defense Acquisition Corps and is DAWIA Level III certified in both Contracting and Information Technology. He earned his bachelor's degree with honors and a master's degree in Business Administration with an Information Systems Management concentration from Old Dominion University and a PhD from Capella University in 2009. His study put forth one of the first quantitatively derived mathematical models for forecasting Enterprise Resource Planning System ROI in the public sector, and was based on data collected from Naval and DLA field activities.

Thomas Wardwell—is a 1986 graduate of the United States Naval Academy and spent 20 years as a Supply Corps Officer honing his skills in logistics, contracting, and information systems. He volunteered as the Head of Contracts supporting contingency and counterterrorism operations in the Horn of Africa, where he envisioned and implemented an electronic contract writing solution to create a data-centric procurement environment with improved data visibility and a central reporting capability to facilitate enhanced decision making. He currently serves as a Navy civil servant improving electronic contract writing capabilities and acquisition business systems and defining the capabilities for a future department-wide procurement system while overseeing the DoN's portfolio of 120 systems and applications supporting the Acquisition Business Mission Area.

Randall Mora—is Founder, President, and CEO of Avum, Inc. In Avum, he has built, and continues to build, a professional services and computer sciences research and development firm that specializes in large-scale systems and software development. With more than 35 years of experience in systems architecture, development, and deployment, Mora is currently focusing on cognitive computing approaches and enterprise architectures to solve complex mission critical solutions for achieving interoperability between cross-domain disparate systems. Mora is currently applying this technology across multiple contracts, including major government and commercial business applications.

Introduction

Frank Kendall, then Under Secretary of Defense for Acquisition, Technology and Logistics, released the first defense acquisition system performance report in June 2013. This report focused primarily on performance related to the collective outcomes of Major Defense Acquisition Programs (MDAPs), but additionally explored various descriptive dimensions and acquisition approaches of the same (Kendall, 2013). Each annual report builds on the work previously conducted, and focuses on data-driven analysis relying on statistical techniques to identify trends that improve the defense acquisition community's insights into how contract incentives are motivating better contractor/vendor performance (Kendal, 2016).

Nevertheless, large amounts of data (in modern jargon, “Big Data”) are now available for research in the area of defense acquisition. Over the past several years, changes in electronic commerce have increased the amounts of both structured and unstructured data available—both in runtime and archived environments. This electronic data, from a variety of different acquisition agencies, can be obtained by a variety of means and used for a multitude of purposes (Snider et al., 2014).

Traditional statistical and trend analysis methods thus far have been primarily relied upon to explore trends and test metrics in the sets of acquisition data at hand. Sometimes, spreadsheets of linear regression correlation are employed, or, in some more modern applications, multivariate structural equation models via scientific applications such as SPSS and AMOS are leveraged for their ability to evaluate complex variable relationships, such as nested or recursive if-then patterns (Byrne, 2016).

However, not only are today’s modern datasets large in magnitude, they are also large in variety and complexity (Gartner, 2013). Furthermore, to address this state of data, new statistical modeling techniques, more powerful than before, have had to be created. This is due to the older methods finding difficulty with some of the size problems Big Data represents, such as privacy and security concerns (Parms, 2017). Thankfully, computer power necessary to employ the modern techniques is less expensive today, the software near free, and the storage capacities available now yield bewildering capacities at a fingertip, and with amazingly fast access speed. In fact, these performance parameters appear to continue along a Moore’s trend line against critical opposition (Magee, Basnet, Funk, & Benson, 2015). Presently, one of the more interesting of the new statistical modeling techniques is *neural network algorithm machine learning*.

Neural network modeling involves utilizing a “powerful computational data model that is able to capture and represent input/output relationships.” This model was developed out of the desire to create artificial intelligence systems capable of completing functions that were previously executed solely by the human brain. One benefit of using neural network modeling lies with its capacity to display and comprehend both linear and non-linear relationships from the data to which it is supplied (NeuroSolutions, 2015).

Research Question

Because “Big Data” is present in the Defense Acquisition Business space, and, because the demand to critically understand real cause-and-effect relationships between variables within that data is persistent from the Acquisition community, this paper’s research question is, *Can a neural network modeling technique be confidently relied upon to meaningfully explore variable relationships within acquisition business datasets?* Because, if it is, then any question may be reasonably asked by anyone of such a dataset; and, via the neural network-enabled tool, the answers they receive will come with scientific statistical confidence as to whether they can be trusted as interesting or useful answers.[1] In order to explore this research question, the study opted to use business data on contractor performance and attempted to isolate predictive variables from past performance information predictive of good performance.

[1] The role of human judgement of course, notwithstanding.

Methodology

This research uses the *Simple Action Research Mode*l (MacIsaac, 1995). Direct participation by the researchers went into answering the research question. In accordance with the steps of direct action: (1) the problem was defined (i.e., Can neural network modeling be applied to Big Data sets in acquisition?); (2) an Action Plan was developed (described in detail below, but generally it was to obtain a subset of Big Data, cleanse it for use, program a neural network tool, write hypotheses postulating expected correlative relationships between variables or variable sets, and execute testing of the hypotheses via the neural network for validation); (3) Execution of the Plan (which was a success: the data was obtained and cleaned, the hypotheses generated, and the neural network tool coded, tested and exercised over several cycles); and (4) Learning and Evaluation, which was completed via the documentation of results in this paper.

It is important to note that the Simple Action methodology employed here is evaluating the paper's research question regarding the applicability of the neural networking modeling technique to big data in the acquisition environment; as such, the actual statistical correlative output of the hypothesis are of a secondary value only (i.e., they are for the purpose of experimenting with the neural network environment itself, as opposed to for discovery in their own right).

Creation of the Cognitive Learning Environment

To build the data environment for evaluation, multiple sources of acquisition data were imported and fused together with iterative slices of multiple groups taken out for analysis. Further breakdown of the environment is described in the Study Plan section. Multiple open-source data analysis and machine learning tools were used to iteratively create models and generate graphs of the data slices. Human evaluation was involved in looking for patterns in the data which may have explained best performance across programs and portfolio groups in the past in order to produce a testable hypothesis. The underlying goal was to find patterns with the best chance of explaining *contractor/vendor performance improvement.*

The cognitive environment is generated in two phases: a simulation phase and a predictive phase. Simulation phase models generated from simple datasets perform predictive analytics. The predictive phase models were generated via Predictive Model Markup Language (PMML)[2] and integrated into a simple prototype for the proof of concept. PMML was used as a standard to integrate defined and tested models into the decision support toolsets. Once those models were iteratively perfected (i.e., acceptable levels of false positives were observed based on training and testing the datasets), we exported the PMML from the models and integrated the new capability into our decision support components.

---

[2] The Predictive Model Markup Language (PMML) is an XML-based predictive model interchange format conceived by Dr. Robert Lee Grossman, then the director of the National Center for Data Mining at the University of Illinois at Chicago. PMML provides a way for analytic applications to describe and exchange predictive models produced by data mining and machine learning algorithms. It supports common models such as logistic regression and feedforward neural networks (Wikipedia, n.d.).

Building the Cognitive Learning Application Framework depended on a predictive model being *able to learn from past experiences and make significantly intelligent decisions.* Thus, bringing together the archived acquisition data and building a model exportable to PMML was the main concern of this research. Figure 1 outlines the researchers' process of creating the Cognitive Learning Environment: mining, fusing, and modeling the datasets. It is important to note the methodology is iterative in nature, requiring the team to return to previous steps during model development. For example, during the Patterns/Analytics step, the need to slice the data differently was identified, which necessitated new data, cleaning, transformation, and so forth. Also, through the development of reusable components, time required for iterations was significantly reduced. Normally, Data Selection and Preprocessing are the most time-consuming steps; usually taking around 80% of the total effort required to build an analytical model (Baesens, 2014).

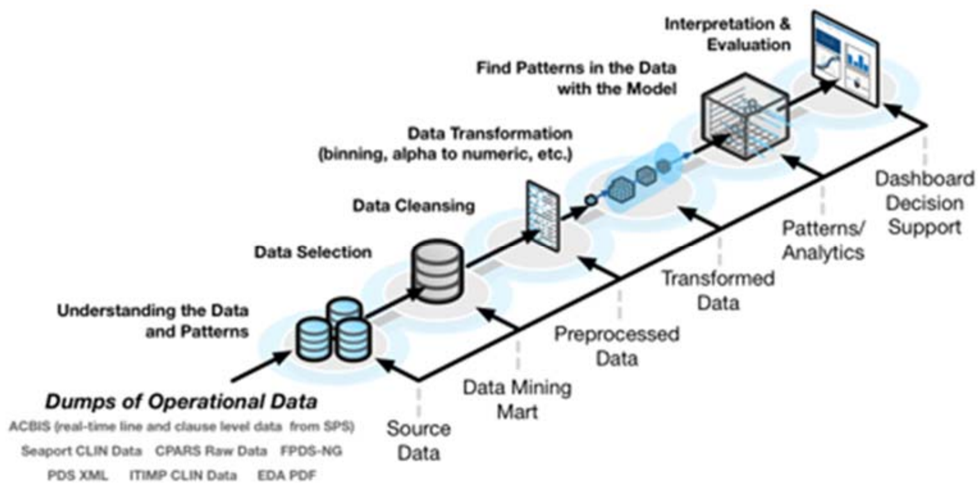

Figure 1. Project Methodology

Hypotheses

This research observes it is commonly claimed: contracts structured with incentivized performance line items (i.e., Cost Plus Incentive Fee, Cost Plus Award Fee, Fixed Price Incentive Fee, Fixed Price Award Fee, Cost Plus Fixed Fee, etc.) are associated (or expected) to enjoy good or better performance than otherwise structured; shorter duration contracts perform better than longer duration ones; competed contracts perform better than sole-source awards; negotiated clauses have an impact on performance, either for good or ill.

Recently, a vendor measure of performance has become available for calculation: the Superior Supplier Incentive Program (SSIP) composite score.

Therefore, this paper takes as its set of testing hypotheses, for the purpose of direct action experimentation of a neural network environment toolset, the following:

- H_1—Contract structures incentivizing performance (i.e., CPIF, CPFF, CPAF, FPIF, FPAF) result in higher Superior Supplier Incentive Program (SSIP) composite scores.
- H_2—Shorter contract duration results in better performance outcomes.
- H_3—Contracts that are competed result in better performance outcomes than sole-source.

- H_4—The mixture of negotiated clause inclusions impacts vendor performance outcomes.

Data Collection

This research began with utilizing an SSIP sample set of contracts from the FY16 review (i.e., FY13, FY14, and FY15 contracts). This data represents the study source boundary. The SSIP master data was derived from the Tri-Service SSIP Selection Methodology inclusive of subsets of CPARS[3] master data by suppliers[4], thus, initially limiting the dataset for analysis (Wardwell et al., 2016).

Analysis proceeded grouping data by similar Product Service Code (PSC) portfolios and similar dollar ranges. Performance was indicated by composite SSIP score (ranging from 0 to 4). Varying contract structures within the SSIP sample set were investigated for potential correlations between contract-type, CLIN mix, contract length, extent-competed, and clause inclusions. Subsequently, these relationships were analyzed within vendors or specific programs and contrasted against available program metrics aside from CPARS metrics.

Core data inputs were derived or pulled from acquisition data sources to which DASN (AP) had access. The following is the comprehensive list of data sources used throughout this research:

- SSIP sample set of contracts from the FY16 Review
- Army Contracting Business Intelligence System (ACBIS)
- Standard Procurement System (SPS)
- FPDS-NG[5]
- PDS XML[6]
- ITIMP CLIN Data[7]
- EDA PDF[8]

Initially, the SSIP sample set of contracts from the FY16 review matched corresponding ACBIS contracts. This allowed CLIN-level information from SPS and DFARS clause inclusions to be included for analysis. Table 1 outlines each attribute and the data source from which they were sourced. Some or all of the attributes were used during different phases of the research. As an example, during the beginning of the project, datasets were limited to the SSIP FY16 Review boundary. This limited research to 1,762

---

[3] Contractor Performance Assessment Reports System

[4] To determine which companies and business segments will be rated in a given FY, the Services use USASpending.gov to aggregate Systems contracts' obligations for the last three FYs by supplier and business segment. For each agency, the funding agency can be found by using the funding codes DoN (1700), AF (5700), and ARMY (2100). The obligations are maintained for all companies in the Air Force Industrial Liaison Office data warehouse. However, only the Top 100 or so (by obligation amount) are pulled for SSIP consideration.

[5] Federal Procurement Data System-Next Generation

[6] DoD Procurement Data Standard, Extensible Markup Language

[7] Navy's Integrated Technical Item Management and Procurement system, Contract Line Item

[8] DoD Electronic Document Access system, Portable Document Format file

records. After cleansing and transformation (i.e., matching/fusion) 972 contracts remained. Analysis of this dataset led to a determination that using source attributes directly from CPARS and including the full set of CPARS data (e.g., 174,138 records), which included contracts from the Navy, the Army, and the Air Force, would be a preferable set for testing the models.

Later iterations of the dataset focused on contract-level details and CPARS ratings. SPS line item details from ACBIS were suspended.

Table 1. Data Attribute Sources

Attribute	Sourced From
Contract-Type (CLIN mix)	ACBIS (SPS line item details)
Contract-Type (contract-level)	FPDS-NG
Extent-Competed	FPDS-NG
Contract Length	FPDS-NG/CPARS
DFARS Clause Inclusions	ACBIS
CLIN Count	ACBIS (SPS line item details)
CPARS Award Value	CPARS
Quality, Schedule, Cost Control, Management, Small Business	CPARS
SSIP	SSIP from the FY16 review
PSC & Portfolio Group	Defense Procurement and Acquisition Policy (DPAP) office

Data Selection, Cleansing, and Transformation

Data selection, cleansing, and transformation of this research was continuous throughout, and encompassed a significant amount of the effort in bringing operational data together correctly. Initial assumptions changed based on accuracy of data, and cleaning what became required. For instance, CPARS data is notoriously inconsistent when it comes to referencing IDIQ contracts. Sometimes, the CPARS references the base IDIQ, which is meaningless for analysis purposes. In other instances, the task order number is referenced.

The following outlines tasks performed to conduct preliminary data selection, grouping, and fusion between sets with the intent of matching with fusion while mitigating issues preventing a dataset:

1. Load of individual data into a data repository for selection and cleansing.
2. Determination of which award-value (CPARS) or obligation-amount (FPDS-NG) would better serve the research objective.
3. Analysis of the FPDS-NG contract numbers in preparation for matching to CPARS data. Contract number formatting varies across different source data repositories, thus, care must be taken to ensure data across those sources match properly. For instance, non-alphanumeric characters (dashes, spaces, etc.) must be removed, and the first 13 characters must be selected, unless the contract number starts with "GS," "HHS," "LC," "NN," or "V." In those cases, the last 13 characters are selected. This method yields a 94% match

with FPDS-NG PIID[9] or Referenced IDV[10] PIID; however, matching algorithms are modified and tuned until the highest percentage match across the data emerges.

4. Contract-type, extent-competed, contract-length and PSC were selected from the base contract or mod with the highest obligated amount change.
5. For indefinite delivery/indefinite quantity (IDIQ) contracts with "D" in the 9th position of the contract number, the algorithm took into account both the "contract-num" and "order-num" values for successful match. Order number was parsed from the CPARS contract number and used both contract number and order number to match with FPDS-NG data using the highest obligated amount change of the mods (28% of the CPARS contracts). IDIQ contracts without order numbers were removed from the CPARS data (27%), as these would not produce meaningful results.
6. Portfolio and Portfolio Groups for each contract were identified by matching the FPDS-NG PSC code with data from the DPAP Product and Service Code (PSC) Selection Tool.

Analytics & Modeling

To perform the analytics and modeling of the data itself, different representations were built of results (e.g., Excel pivots, scatter plots, graphs, charts, etc.). This first step was conducted to steer further neural modeling efforts. Working with defense acquisition subject matter experts, iterative evolutions of the dataset were visualized to attempt discovery of patterns representing algorithms executable at high levels of accuracy. The iterations generated numerous views into the dataset, and participants were successful in coming to an understanding of the best way to begin modeling. Visual data exploration proved important in supplying initial insights into the data, which researchers then adopted throughout the modeling (Baesens, 2014).

To be useful to the community at large, end users should not have to acquire and learn complex analytic software to obtain predictions from these models. To facilitate this goal, once models are trained and tested, they're anticipated to be exported as PMML. The PMML can then be fed into a wide variety of systems and programming languages, which can be used to run the model(s) against incoming data. A Gartner 2017 Magic Quadrant for Data Science Platforms (called in 2016 "Advanced Analytics Platforms") evaluated a new set of 16 analytic and data science firms over 15 criteria and placed them in four quadrants, based on completeness of vision and ability to execute (Piatetsky, 2017). KNIME[11] was in the top quadrant with SAS, IBM, and RapidMiner. This study found KNIME to exhibit a flexible and extensible design, display an ease of use and verified capability to export PMML (essential to support of developing this research's data modeling and Cognitive Learning Application Framework).

The KNIME platform was selected for this study's use as the neural network tool. The KNIME advanced analytics platform is an open-source analytics platform, and was used

[9] Procurement Instrument Identifier
[10] Indefinite Delivery Vehicle
[11] KNIME Analytics platform: https://www.knime.org/products

to run computational analysis against the historical dataset to generate various sets of predictive models. Numerous models, such as polynomial and logistical regressions, were used to predict numerical values for variables, whereas *neural networks* and *decision trees* were used to predict categorical values for variables. Operational data was loaded into local databases (i.e., PostgreSQL) and quickly iterated regardless of it being functional in Excel for preliminary analysis or in KNIME for modeling.

Figure 2 illustrates the modeling process for the paper's problem set. Significant time was spent segmenting the dataset and running the data through models until patterns emerged that are relative to the expected results sought (i.e., contract structures performance).

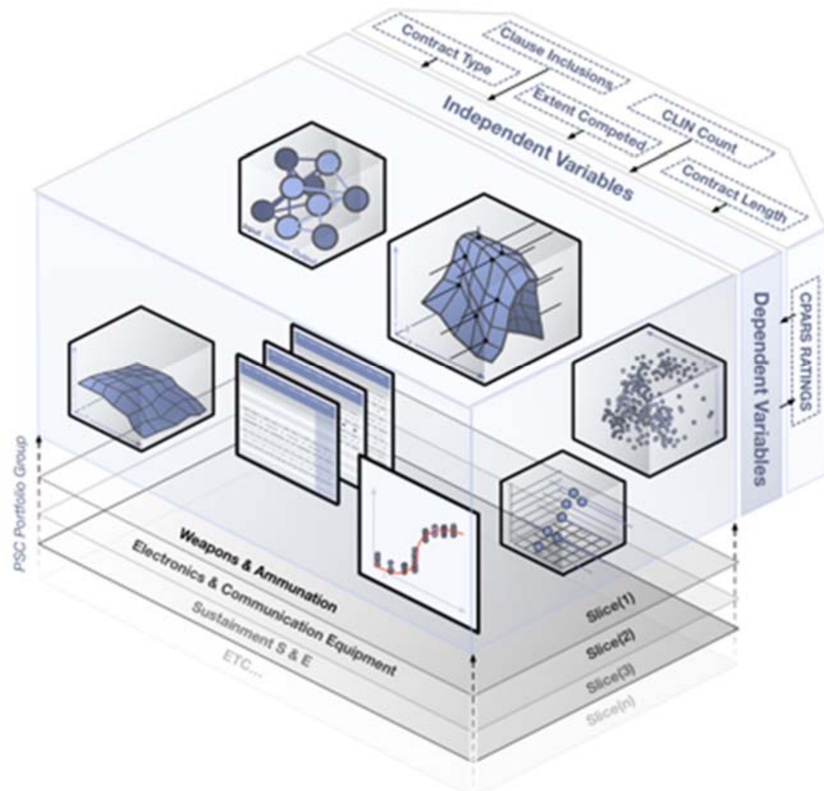

Figure 2. Analytics and Modeling Paradigm

Figure 2's illustration reveals how data in the study was segmented by PSC Portfolio Groups and used CPARS ratings (both by composite SSIP scores and individual ratings values) as dependent variables. In contrast, numerous independent variables were inputs to the modeling process, only some of which are depicted.

Data Analysis and Results

Modeling used contract-type and extent-competed as independent variables with composite CPARS ratings as dependent variables. The research goal at this stage was to find patterns in the data revealing correlations between contract-type and extent-competed to be neural network modeled, and, thus, exercise the paper's hypotheses via statistical correlation testing and address the paper's research question. Contract-length became another independent variable added to the set for modeling and, from that point forward, any attribute available and deemed pertinent to the research hypothesis objective was incorporated into operational data or transformed and sanitized. As a result, multiple back and forth passages through previous steps of data selection, cleansing, and transformation were avoided. Excel pivots were created against the full CPARS dataset, which supplied 174,138 records for final analysis in the study. No obvious significant pattern emerged from observing the pivot tables. Table 2 summarizes initial descriptive data for contract-type in the sample.

Table 2. Pivot of Contract-Type and SSIP[12]

Contract Type Description	Average of ssip	Count of ssip
COMBO	3.17	3,617
COST NO FEE	3.08	9,047
COST PLUS AWARD FEE	3.22	4,801
COST PLUS FIXED FEE	3.05	31,491
COST PLUS INCENTIVE FEE	2.76	3,380
COST SHARING	2.42	17
FIRM FIXED PRICE	2.82	116,660
FIXED PRICE AWARD FEE	3.12	366
FIXED PRICE INCENTIVE FEE	2.76	1,902
FIXED PRICE LEVEL OF EFFORT	2.86	522
FIXED PRICE REDETERMINATION	2.87	175
FIXED PRICE WITH ECONOMIC PRICE ADJUSTMENT	2.84	714
ORDER DEPENDENT	2.43	209
Grand Total	**2.89**	**172,901**

Table 3 summarizes initial descriptive data for extent competed in the sample.

Table 3. Pivot of Extent-Competed and SSIP[13]

Extent_Competed	Average of ssip	Count of ssip
COMPETED UNDER SAP	2.71	4,979
FOLLOW ON TO COMPETED ACTION	3.03	260
FULL AND OPEN COMPETITION	2.93	78,917
FULL AND OPEN COMPETITION AFTER EXCLUSION OF SOURCES	2.88	35,975
NOT AVAILABLE FOR COMPETITION	2.98	20,036
NOT COMPETED	2.80	31,516
NOT COMPETED UNDER SAP	2.87	2,150
Grand Total	**2.90**	**173,833**

Working With the Predictive Analytic Environment

The pivot table input yielded prima facie evidence against H_1 (that incentivized contract-types yield greater contractor performance) and H_3 (that competed contracts yield greater contract performance). With this starting point in hand, the study moved into the regression analysis phase and began running polynomial and linear regression models

[12] 1,237 unmatched records from Table 2 were removed

[13] 305 unmatched records from Table 3 were removed.

against the data with KNIME software to evaluate how close target measures of interest were with hypothesized outcomes. We ran both linear and polynomial regression algorithms against the data to attempt to find a mathematical formula that would take the extent-competed and contract-type as inputs and give a predicted value of the SSIP score.

Notes on Inputs: The regressions require numeric values for the dependent and independent variables. The SSIP score (the dependent variable) is a numeric value; however, the independent variables (contract-type and extent-competed) are strings. The type was converted to numbers by assigning a unique integer value to each distinct value (i.e., firmed fixed price was set to 10).

Figure 3 illustrates the KNIME's workspace used for these regressions. Node 1 utilizes an Excel file as input, and the data is passed into a partitioning node (Node 6) to give us a set of data on which to run the regressions, and a small set of data used to test the regressions. The data is passed to three different nodes. Node 2 runs the polynomial regression (with a max polynomial degree of 2), Node 10 runs a linear regression and Node 11 generates a 3D scatter plot to assist in visualizing the data.

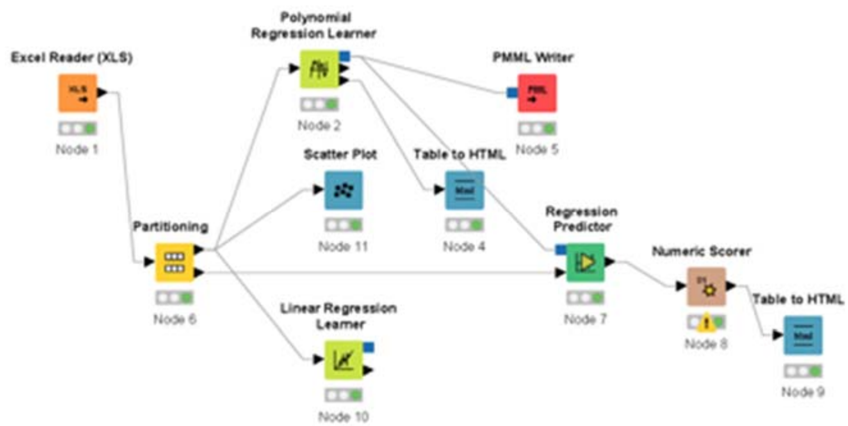

Figure 3. Initial Polynomial and Linear Regression

The results of the linear regression analysis showed no linear correlation existing in the dataset between independent variable types of "extent-competed" or "contract-type" and dependent variable types of contractor SSIP score. The results of the Polynomial Regression were pushed into a Regression Predictor and Numeric Scorer to output the model's statistics. Figure 4 shows all learned coefficients for the contract-type and extent-competed.

Regression involves numerous variables. The R^2 statistic, listed at the bottom of the figure, measures the proportion of variability in the SSIP score and CPARS rating (dependent variables) that can be explained by the contract-type and extent-competed (independent variables). This value is between 0 and 1, and can be negative, with a value close to 1 indicating that a large proportion of the variability in the response can be explained by the regression. The standard error is the average amount that the response will deviate from the true regression line (James et al., 2013).

Statistics on Polynomial Regression

Variable	Coeff.	Std. Err.	t-value	P>|t|
Contract_type_numeric	0.0528	0.0501	1.0537	0.2922
Extent_competed_numeric	0.0682	0.0806	0.8459	0.3977
Contract_type_numeric^2	-0.0114	0.0074	-1.5413	0.1235
Extent_competed_numeric^2	-0.0216	0.0177	-1.2241	0.2211
Intercept	2.4667	0.1129	21.8573	0.0

Multiple R-Squared: 0.4437
Adjusted R-Squared: 0.0047

Figure 4. Statistics on Polynomial Regression

The Numeric Score (Node 8) in the flow computes certain statistics between the numeric column's values (r_i) and predicted (p_i) values. It computes $R^2=1-SS_{res}/SS_{tot}=1-\Sigma(p_i-r_i)^2/\Sigma(r_i-1/n*\Sigma r_i)^2$ (can be negative!) ("Coefficient of Determination," n.d.); mean absolute error ($1/n*\Sigma|p_i-r_i|$) ("Mean Absolute Error," n.d.); mean squared error ($1/n*\Sigma(p_i-r_i)^2$) ("Residual Sum of Squares," n.d.); root mean squared error ($sqrt(1/n*\Sigma(p_i-r_i)^2)$) ("Root-Mean-Square Deviation," n.d.); and mean signed difference ($1/n*\Sigma(p_i-r_i)$) ("Mean Signed Deviation," n.d.). The computed values can be inspected in the node's view and/or further processed using the output table. Table 4 contracts the results from the Numeric Scorer.

Table 4. Numeric Score Results

Prediction (ssip_raw)	
R^2	0.019
Mean absolute error	0.55
Mean squared error	0.458
Root mean squared error	0.676
Mean signed difference	-0.018

Next, we built more complicated predictive analysis algorithms in an attempt to find the correlations/predictive capabilities we were expecting (i.e., model contract structures that incentivize performance correlated to better performance outcomes). Figure 5 shows the training and testing of two different types of neural networks. It initially splits the data into a training set, and a validation set: we train the neural network on the training set and measure the performance on the validation set. Node 5 partitions the data (i.e., 80% Training and 20% Validation) and pushes the sets into both an RProp MLP Learner and a PNN Learner (DDA).

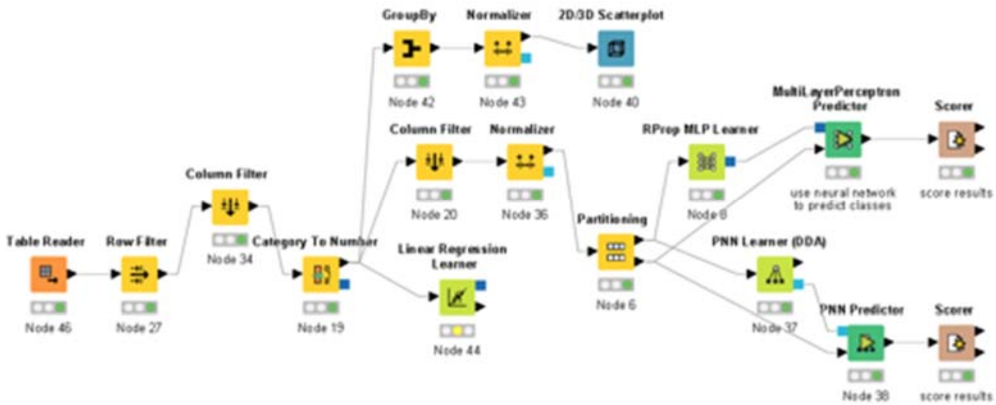

Figure 5. Neural Network Model

The RProp algorithm is used for multilayer feedforward networks. RProp performs a local adaptation of the weight-updates according to the behavior of the error function (Riedmiller & Braun, 1993).

The PNN Learner (DDA) trains a Probabilistic Neural Network (PNN) based on the DDA (Dynamic Decay Adjustment) method on labeled data using Constructive Training of Probabilistic Neural Networks (Berthold & Diamond, 1998) as the underlying algorithm. This algorithm generates rules based on numeric data. Each rule is defined as a high-dimensional Gaussian function that is adjusted by two thresholds, theta minus and theta plus, to avoid conflicts with rules of different classes. Each Gaussian function is defined by a center vector (from the first covered instance) and a standard deviation, which is adjusted during training to cover only non-conflicting instances. The selected numeric columns of the input data are used as input data for training, and additional columns are used as classification targets. Either one column holding the class information, or a number of numeric columns with class degrees between 0 and 1, can be selected. The data output contains the rules after execution, along with a number of rule measurements. The model output port contains the PNN model, which can be used for prediction in the PNN Predictor node.

The KNIME tool was able to create an abstraction layer around the algorithms that were being run in the neural network nodes. A basic understanding of what the nodes are doing and how to configure them for optimum performance facilitates quicker iterations and an overall feasibility study that supports rapid prototyping.

The MultiLayerPerceptron and the PNN Predictor nodes are used to validate the resulting trained model against the test dataset. For the MultiLayerPerceptron Predictor, if the output variable is nominal, the output of each neuron and the class of the winner neuron are produced. The PNN Predictor is doing a similar test (i.e., using the trained model to validate with the test data). In this case, it also outputs predicted data with an additional classification column (e.g., CPARS quality attribute).

Figure 6 and Figure 7 outline the confusion matrix and accuracy statistics for the model. An example of what it's showing is as follows: For the MultiLayerPerceptron model/predictor, using 20% of the data to score/validate based on contract-type and extent-competed, we successfully predicted 271 times Technical/Quality of Product or Service would be rated as Satisfactory. As well, 166 times Technical/Quality of Product or Service was rated as Satisfactory when it should have been Very Good. We only trained the model to predict Quality at this point. The accuracy for the MultiLayerPerceptron was 45.427% and

48.018% for the PNN Predictor. Evaluation ratings are defined as: E: Exceptional, V: Very Good, S: Satisfactory, M: Marginal, U: Unsatisfactory, and N: N/A.

Confusion Matrix - 0:16 - Scorer (score results)

File Hilite

quality \ Pr...	V	S	E	U	N	M
V	25	166	0	0	0	0
S	10	271	0	0	0	0
E	22	106	2	0	0	0
U	0	1	0	0	0	0
N	2	27	0	0	0	0
M	1	23	0	0	0	0

Correct classified: 298 Wrong classified: 358

Accuracy: 45.427 % Error: 54.573 %

Cohen's kappa (κ) 0.067

Figure 6. MultiLayerPerceptron Predictor

Confusion Matrix - 0:39 - Scorer (score results)

File Hilite

quality \ Pr...	U	V	S	E	N	M
U	0	0	1	0	0	0
V	0	14	164	13	0	0
S	0	8	267	6	0	0
E	0	9	87	34	0	0
N	0	2	27	0	0	0
M	0	1	22	1	0	0

Correct classified: 315 Wrong classified: 341

Accuracy: 48.018 % Error: 51.982 %

Cohen's kappa (κ) 0.13

Figure 7. PNN Predictor

Work on finding correlations within the data continued throughout the research. Through additional tuning and new independent variables, the study found the model accuracy statistics showed positive results. New problem sets with higher accuracy statistics should find themselves into a defense procurement toolset soon and/or a sequel to this research.

PMML Integrated Decision Support Tool Kit

The Java PMML API (GitHub, 2017) is an open-source project that provides a PMML producer and consumer libraries for the Java/JVM platform. Using these libraries, a Java wrapper is created around each model. These wrappers each take a pre-defined set of inputs and provide a predicted output value. These wrappers then become pluggable components that can be added to any Java-based tool, such as websites, web services and stand-alone applications. For example, a tool that runs Pre-Validations on PDS XML[14] (Defense Procurement and Acquisition Policy, 2008) files can quickly be updated to take the information it parses from the PDS file, and feed it to these various models to augment the validation results to include predictive analytics.

Early in the research project, a simple model prototype was built to export the PMML from this model. This PMML was used in a Java application as a Proof of Concept (POC) for the ability to use a PMML-defined model and integrated its predictive capabilities into the Cognitive Learning Application Framework, thus, having an end-to-end solution for decision support capabilities in the application toolsets.

Future work in this space involves making the models used by these wrappers dynamic. The current iteration utilizes static PMML that was generated using a historical dataset. It would be desirable to have the ability to update the models in real time as new data comes in.

Conclusion

This study's research question was, *Can a neural network modeling technique be confidently relied upon to meaningfully explore variable relationships within acquisition business datasets?* This paper's result was positive to the research question.

The study's open architecture framework (i.e., the Cognitive Learning Application Framework [CLAF]) for Acquisition Decision Support and Business Intelligence successfully integrated and prototyped a neural network model using a PMML standard and explored variable relationships using four test hypotheses addressing contract performance data. Regarding the study's test hypotheses, results were inconclusive. Only H_1 (incentivized contract types correlate with higher vendor performance scores) and H_3 (competed contracts correlate with higher vendor performance scores) were thoroughly evaluated, and proved to be inconclusive via initial standard regression technique. Due to datasets being too small for substantive use in big data network evaluation, or, because of time limitations preventing necessary dataset concatenation, H_2 (shorter duration contracts correlate with higher vendor performance scores) and H_4 (contract clauses have impact on vendor performance score) could not be evaluated.

The study's main function went on to explore H_1 and H_3 as a means of addressing the research question via a direct action methodology of research. Experimenting with the neural network mode of analysis, the study attempted accurate prediction of vendor performance scores given an input of one of the hypothesized independent variables. The study's neural network obtained a maximum accuracy score of 49%. Obtaining this level of accuracy required careful, and sometimes tedious, assembly of statistical and logical

[14] The Procurement Data Standard (PDS) is a system-agnostic data standard that is adopted and implemented DoD-wide for creation, translation, processing, and sharing of procurement actions. It defines the minimum requirements for contract writing system output to improve visibility and accuracy of contract-related data, to support interoperability of DoD acquisition systems and to standardize and streamline the procure-to-pay business process. Further, the PDS will improve visibility of contract-related data, enabling senior DoD leadership to make better informed business decisions. And finally, this data standard will support future migration to enterprise and federal systems and processes where appropriate.

components. Although work has stopped, it is anticipated to resume post-publication in order for more potential latent variable relationships to be discovered, or, presumed relationships tested for the potential to be dispelled.

The value of discovering through this study's experience that there's evidence that the neural network modeling technique is applicable to big data sets in acquisition is that, now, any question for which there is a discrete data element available, or derivable, within those sets, can expect a trusted answer (if interesting and useful) with scientific statistical confidence.

Findings

The KNIME advanced analytics platform can write to PMML simply by dragging the PMML writer to the model. Figure 3 illustrates this capability in Node 5 (PMML Writer). This and the Application Frameworks created during this research project will enable future researchers to quickly bring new data into the modeling process, as well as integrate exported PMML models into Defense Acquisition Decision Support tools.

Cognitive computing (neural network modeling) solutions promise better-informed buying and increased compliance, and may make this faster and easier to accomplish once generated. Also, from this experience, the study found that loading more than just enough data helps significantly during the modeling phase.

Given the success of this research, it is recommended that government and industry oversight entities build a cognitive learning component for acquisition support that uses archived acquisition data from known repositories. The component can then be used as a stand-alone tool for the acquisition community and/or integrate into existing acquisition community toolsets and contract writing applications.

Afterwards, the models could be leveraged in decision support or Business Intelligence (BI) dashboards by the acquisition community.

In fact, a simulation tool could even be envisioned that would allow contracting officers (KOs) to perform scenario testing surrounding new agreements—one that would, given any variables, project performance indices based on purchase type or agreement structure, or other discovered latent relationships.

Areas for Further Research

The combination of the Cognitive Learning Acquisition Framework and a Big Data archive together form a methodology for an Application Framework, enabling a dynamic information analysis space to build intelligence into acquisition decision support tools. Future practical research into the feasibility and capability of applications leveraging such a framework could include the following: decision support for contracting officers and program managers surrounding contract structure tied to true historic performance and delivery outcomes; modeling of contract incentives, structures, and policies, and their impact on performance, delivery, costs, and schedule across major programs; volume of modifications and manual and/or late payments tied to contract, clause, and line item structures, as well as the overall quality of the contract data and compliance with contracting rules and regulations; EVM[15] outcomes correlated to initial negotiated contract terms; award and incentive fee payouts tied to EVM and CPARS metrics; vendor past performance within specific product service codes correlated to historic contract structure; and Q&A support for initial acquisition planning (i.e., “What vendors typically support this product or service?”, “What type of contract is most widely used?”, “Is this work typically competed?”, and “What clauses above and beyond the typical prescriptions accompany this type of buy?”). Finally, the incorporation of additional public and Defense datasets within the financial, logistics, and commercial spaces into the Big Data archive is warranted to provide opportunities for further exploration of data relationships for use in acquisition decision making.

References

Baesens, B. (2014). *Analytics in a big data world*. New York, NY: Wiley.

Berthold, M., & Diamond, J. (1998). Constructive training of probabilistic neural networks. *Neurocomputing*. Retrieved from http://www.sciencedirect.com/science/article/pii/S0925231297000635.

Byrne, B. M. (2016). *Structural equation modeling with Amos: Basic concepts, applications, and programming*. New York, NY: Routledge, Taylor & Francis Group.

Coefficient of determination. (n.d.). In *Wikipedia*. Retrieved from https://en.wikipedia.org/wiki/Coefficient_of_determination

Defense Procurement and Acquisition Policy. (2008). Procurement Data Standard and other enterprise initiatives. Retrieved from http://www.acq.osd.mil/dpap/pdi/eb/procurement_data_standard.html

Gartner. (2013). Gartner identifies top technology trends impacting information infrastructure in 2013. Retrieved from http://www.gartner.com/newsroom/id/2359715

GitHub. (2017). Java PMML API. Retrieved from https://github.com/jpmml

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). Unsupervised learning. In G. James, D. Witten, T. Hastie, & R. Tibshirani, *An introduction to statistical learning with applications in R* (pp. 373–374). New York, NY: Springer Science+Business Media.

Kendall, F. (2013). *Performance of the defense acquisition system, 2013*. Washington, DC: Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics.

Kendall, F. (2016). *Performance of the Defense Acquisition System: 2016 annual report*. Washington, DC: DoD.

Magee, Basnet, Funk, & Benson. (2015). Quantitative empirical trends in technical performance. *Technological Forecasting and Social Change.*

Mean absolute error. (n.d.). In *Wikipedia*. Retrieved from https://en.wikipedia.org/wiki/Mean_absolute_error

Mean signed deviation. (n.d.). In *Wikipedia*. Retrieved from https://en.wikipedia.org/wiki/Mean_signed_difference

NeuroSolutions. (2015). What are neural networks & predictive data analytics? Retrieved from http://www.neurosolutions.com/products/ns/whatisNN.html

[15] Earned Value Management

Parms, J. (2017). Privacy and security issues in the age of big data. Retrieved from https://www.business.com/articles/privacy-and-security-issues-in-the-age-of-big-data/

Piatetsky, G. (2017). Gartner 2017 magic quadrant for data science platforms: Gainers and losers. Retrieved from http://www.kdnuggets.com/2017/02/gartner-2017-mq-data-science-platforms-gainers-losers.html

Residual sum of squares. (n.d.). In *Wikipedia*. Retrieved from https://en.wikipedia.org/wiki/Residual_sum_of_squares

Riedmiller, M., & Braun, H. (1993). A direct adaptive method for faster backpropagation learning: The RPROP algorithm. In *Proceedings of IEEE International Conference on Neural Networks* (pp. 586–591).

Root-mean-square deviation. (n.d.). In *Wikipedia*. Retrieved from https://en.wikipedia.org/wiki/Root-mean-square_deviation

Snider, K. F., Rendon, R. G., Soeters, J., Shields, P. M., & Rietjens, S. J. H. (2014). Retrieving what's already there—Archival data for research in defense acquisition. In *Routledge handbook of research methods in military studies* (pp. 78–91). London, England: Routledge/Taylor & Francis Group.

Wardwell, T., Nangia, S., Simeoni, C. J., Brittenham, R., & Campbell, T. (2016). *Tri-service SSIP selection methodology.*

Acknowledgments

We gratefully acknowledge the contributions of numerous individuals who provided assistance and support in this research project. Larry Wright, David Harris, Kevin Chiu, and Anita Doyle each provided significant time in addition to performing their primary work in order to provide assistance in conducting the data loading, cleansing, transformation, and analysis for this project. David Harris's previous training and research with cognitive engineering proved invaluable in bringing everything together, from building the models to integrating the PMML into a Java application.

Special thanks go to David Michalczuk for the outstanding 3-D diagrams he produced for this project. We all love your diagrams.

Finally, the close partnership with DASN (AP) contributed in many diverse ways, including accessing cleansed data that normally is not available to researchers. This made the normally difficult and time-consuming task of acquiring the operational acquisition data analyzed in this project conclude significantly faster than otherwise possible. Their subject matter expertise (SMEs) in acquisition systems, policies, and procedures were immeasurable in our success.

The responsibility for the contents of this study rests solely with the authors. Any errors or omissions are their responsibility and should in no way reflect on any of the individuals named above.

Acquisition Research Program
Graduate School of Business & Public Policy
Naval Postgraduate School
555 Dyer Road, Ingersoll Hall
Monterey, CA 93943

www.acquisitionresearch.net