INDOT: Building the Bridge from Spreadsheets to Machine Learning for Highway Construction Contract Bundling

Brett Boston
Vern Herr
Ken Mocabee

Executive Summary

Machine Learning (ML) is a proven technology for identifying hidden data relationships and patterns. To our knowledge, it had not been previously utilized for transportation project bundling. While extensive research by Purdue University, KPMG, and State Agencies has documented that construction costs will be reduced with project bundling, practical application has proven complex, time-consuming, and subjective. The application of Machine Learning's sophisticated analytical tools enable contract bundling staff to leverage the full potential bundling offers.

This paper reviews the findings from a Machine Learning research project that bundled over 3,400 transportation construction projects in Indiana. Further, Machine Learning bundled historical packages, previously bundled by INDOT expert asset selection teams. Four years of historical INDOT expert manual bundling were compared to the results generated on the same projects by Machine Learning. The study confirmed the bundling application dramatically improved the ability of planners to identify better project bundles and to generate an additional 20% in potential cost savings. Additionally, INDOT was able to apply consistent state-wide decision criteria across all projects, save almost 1,000 hours of engineering time, and accelerate the start to finish time of the bundling process.

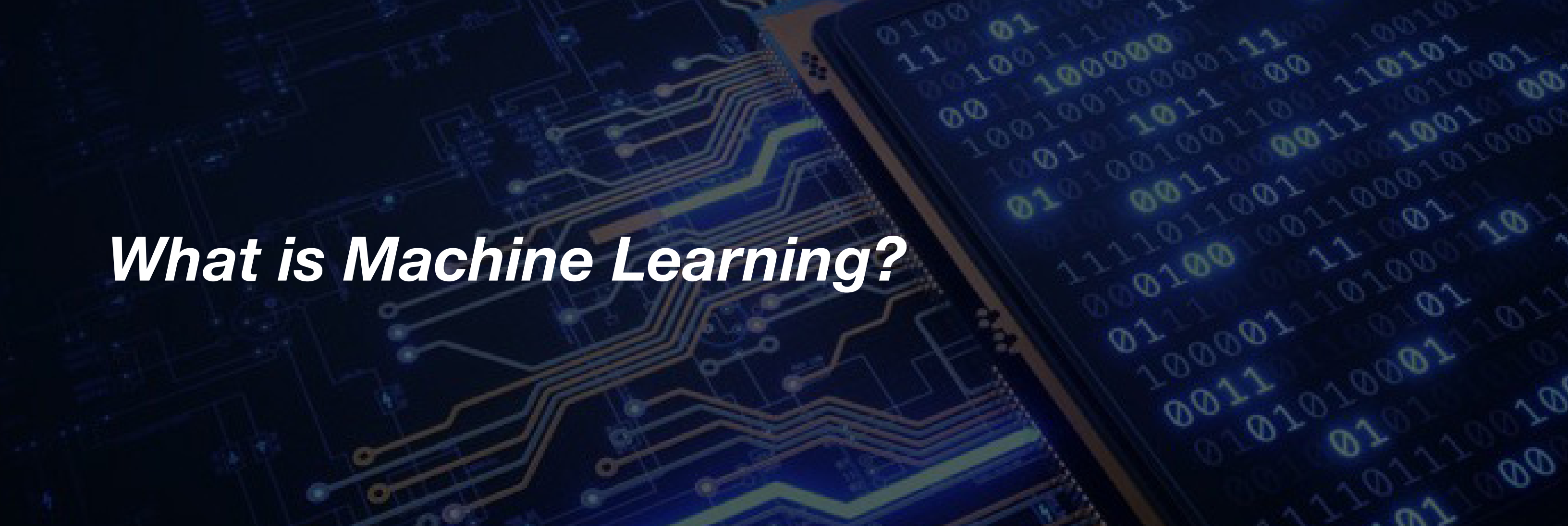
An iterative combination of the Machine Learning followed by expert review enabled:

- Consideration of a wider and more diverse range of project bundling options
- Evaluation of multi-year bundling scenarios
- Faster and better decision-making

Results

- **Comparing identical 2021-2024-year projects, the research project identified $107 million more in projected cost savings than the previous INDOT method**
- **Staff time to identify these bundles was reduced by 60%**
- **Improved corridor bundles will improve traffic flow and reduce construction delays in Indiana**
- **Machine Learning can produce high quality contract bundles from thousands of available projects in minutes, enabling the experts to focus on a more strategic project considerations**

This review also includes some lessons learned at INDOT that may help other agencies considering how to implement ML for cost savings, time reduction and greater operational effectiveness.



What is Machine Learning?

Machine Learning is a unique subfield of artificial intelligence in which algorithms learn to fulfill tasks. Machine Learning algorithms build mathematical models that are based on sample data to make predictions or decisions without being explicitly programmed to do so. The unique design feature of these algorithms is their ability to “learn,” adapt and improve with successive runs.

Overview

Each year, INDOT evaluates over 1,000 programmatic bridge, highway structure, safety, mobility and other specialty work projects as candidates for bundling. How these highway construction projects are scoped, packaged, and delivered can greatly influence bid prices, traffic impacts, and Agency effort during procurement and delivery.

INDOT recognized the potential of Machine Learning to produce a consistent starting point for administering the bundling process. From the outset of this research project INDOT established the following goals:

- **Maintain the integrity of the existing INDOT construction bundling processes**
- **Generate value by implementing logical project groupings**
- **Create objective criteria for evaluating the potential value-add of project groupings**
- **Increase procurement efficiency with fewer and larger bid packages but still maintain a competitive bidding process**
- **Improve project management with fewer contracts to manage**

INDOT leaders hoped Machine Learning would assist with the complexity, project volume, and limitations of the bundling factors that could be evaluated in a manual staff-driven process. Identifying project bundles required multi-day meetings of District and Central Office engineers and other subject matter experts. Working from spreadsheets and other tools, INDOT staff identified logical contract bundles, in complex meetings to review, negotiate and prioritize decisions.

It was a difficult, time-consuming and largely manual process fraught with high levels of subjectivity.

What is a Project Bundle?

INDOT classifies project bundles two ways:

Portfolio Bundle: A Portfolio Bundle is a grouping of projects of the same asset type (e.g. bridge projects, highway structure, safety, mobility projects and other specialty work).

Corridor Bundle: A Corridor Bundle is a grouping of project of the same or different asset types in a continuous work zone. These grouping are especially useful for consolidating multiple projects on a specific highway or interstate over multiple years. The most significant cost savings from the bundling process results from:

- **Bundle size**
- **Homogeneity of bridge types**
- **Homogeneity of work types**
- **Geographic concentration**
- **Similarity of site conditions**
- **Flexibility of contract scheduling and sequencing**
- **Faster delivery**
- **Increased productivity of labor and equipment**
- **Complimentary and cost-effective Management of Traffic (MOT) across bigger projects**

Algorithm Development

The bundling application development process began with an in-depth consultation that documented INDOT's formal business rules for manual project bundling. A data dictionary of project data fields, business rules and decision criteria were developed for training the Machine Learning. The Machine Learning algorithm was applied to data from FY 21 through FY24 to produce bundles. INDOT had already created manual bundles for these years. This provided a test of the Machine Learning's ability to replicate decisions made by the INDOT team. The initial goal was to deliver an algorithm that could identify and replicate at least 80% of the bundles selected by INDOT Engineers.

The following definitions were agreed to:

Project – A specific set of highway construction work

Bundle – A collection of projects that have been bundled by the application program – The complete set of data that includes all bundles and projects Parameter – A variable that changes the output of the algorithm

Data Set – a set of project data to be imported into SPMS – The INDOT system used to manage projects

Two important, but unstated needs emerged in this discovery phase. For specialty bridge projects, the ability to group many small contracts under a single prime contractor for improved project management of subcontractors was a known, but informal business rule. The need to manage project bundles across union locals and non-union counties was another business rule that was applied in the manual process and needed to be replicated.

Initial Approach to Bundling

INDOT uses a popular asset management software product, dTIMS for inventorying and managing highway construction products and predicting asset needs. This product is used in approximately 35 U.S. State DOTs. Each of the projects in the INDOT dTIMS database was assigned the following data elements

- Estimated Project Cost
- Latitude/Longitude of each project's midpoint
- Corridor (Route number)
- County Union/Non-Union affiliation

The initial method for creating project bundles relied on a "brute-force" approach. Projects were run through a vast number of multiple iterations. This was successful, but extremely time-consuming.

Simulated annealing was a more efficient approach. This technique provided options for selecting sets of bundles and thresholds that triggered moving projects in/out of a bundle. These included:

- Identification of the strongest candidates enabling the most obvious candidates to be isolated from other packages
- A semi-transparent soft hold for potential candidates that allows them to be included in future runs. This allows permit ongoing review/reconsideration
- A re-scatter and re-run alternative that starts with each project in its own bundle and randomly moves projects to new bundles
- If the total bundle score is better, the bundle is retained.
- If not, it reverts to the previous bundle.
- The bundle keep algorithm will also weigh the decision to keep bundle. It will be kept more often toward the end of the run. Swaps are more likely at the beginning of a run.

This process was repeated until a pre-defined maximum number of iterations was completed. The algorithm can be set to compare up 10,000 bundles or configured to evaluate up to 20 bundles per project.

Initial Prototyping and Establishing a Scoring System

Approximately 1,000 INDOT projects per year were run through a prototype algorithm that produced predictive project bundles. Initial results were encouraging: in the first round, the process identified 89% of the project bundles identified by INDOT engineers.

A closer review revealed that this level of performance was actually better than it appears as a number of manually produced historic bundles didn't follow INDOT's stated business rules. One of the larger ones, a 92-mile Interstate project showed clearly the effect of inconsistency within the manual process and some inefficiencies. The key point is that the manual process was NOT assumed to be correct but was to serve as a comparator for the Machine Learning data. In fact, many of manual bundles were based on incorrect assumptions and misapplication of the business rules.

The Machine Learning highlighted key assumptions in the manual process that were incorrect and found additional cost savings patterns that INDOT staff's manual rules did not allow for.

Refinements

Based on this initial success, 2022-23-24 projects were analyzed and the algorithm was refined.

The INDOT team offered great feedback on Machine Learning and decision criteria refinements.

As predicted, the predictive accuracy of bundling algorithm improved with successive bundling runs.

One of these refinements was the creation and comparison of multiple "what-if?" scenarios. Users were enabled to set the parameters of the bundle, including being able to adjust and compare project distances, whether the projects needed to be bundled by Unions or specialty project, or linked by corridor route or route number. This made it easier for users to move bundles real time, compare them and optimize scenarios. The big idea is Machine Learning takes minutes to create a bundle scenario but many hours if done manually. The sheer complexity of bundling manually means that generally, a one-pass solution with limited options is all that can be addressed. Machine Learning allowed the what-if scenarios to be generate with a few computer keystrokes.

The savings score of each bundle was calculated and rated on a scale from 1-10 with 10 being the best-rated bundles. Multiple bundle runs were generated and compared. The algorithm was modified so each run could produce a unique set of bundles for evaluation.

Additionally, capabilities were added enabling users to select individual bundles and to analyze multi-year bundle alternatives. This enabled construction of single or multiple bundling scenarios and evaluation for cost savings, time reduction and operational efficiencies.

Additional Refinements

Following these refinements, a dashboard was created to display the number of projects in each bundle, a map that showed the special placement of the projects in the bundle, the maximum distance of the projects from each other, as well as the estimated cost savings and overall savings score. The dashboard displayed savings score and comparison of bundle packages. Once groups of bundles were approved, they could be accepted and added to a final accepted bundle program. Any remaining projects could then be added to existing bundles manually, where that made sense to the engineers, or left as stand alone projects.

The accepted bundle display showed all the accepted bundles across bundle runs, and key statistics on the program including the total number of bundles, the average bundle size, the average savings of each bundle and total cost savings.

Additional application and feedback from the INDOT team guided development of three additional bundling alternatives:

- **Identifying Projects by corridor**
- **Identifying Projects by route number.**
- **Identifying Projects by location**

Dashboard views were refined to include:

- **An overall program view of contract size (number of contracts by site)**
- **Grouping of bundles by price point range: $10M, $5-10M, $2-5M**
- **A view to see all bundles for the entire program (by fiscal year)**
- **A side-by-side view enabling more logical program comparisons**

Even after a series of runs, a small number of projects were not assigned to bundles. A manual process, independent of normal business rules, was developed for INDOT engineers to review and consolidate these projects if needed.

At this point, the INDOT team felt there was no further need to manipulate elements at the project level. They were satisfied the bundling application consistently outperformed the prior method.

Based on this assessment, the following process was outlined:

- INDOT Asset management will provide recommendations to each District
- These recommendations will be reviewed and explored for consensus
- Approved projects from the Districts will be placed in a review/deliberation area
- Asset management will work together with Districts to complete bundling
- Leadership will run the bundling application and review potential cost savings
- The team will review bundles and refine them within the bundling application
- INDOT Engineering will add new bundles until all budgeted funds have been allocated
- Only bundled files will be sent to scoping when this is complete
- Once INDOT Engineering has selected project bundles they will be submitted for approval via a web service to the INDOT SPMS Financial Scoping application

Lessons Learned

1. **Machine Learning enables double-loop, deutero learning by looking backward AND forward.** This process enabled INDOT teams to spot and refine the kinds of project bundles that worked best. Multiple data runs enabled tighter, more customized project bundles to be generated, replicated and refined.

2. **INDOT used the combination of the Machine Learning bundling AND expert review to make subject matter experts smarter, faster, and more productive.** The power of this approach lies in the integration of the strengths of Machine Learning to quickly generate bundles and the staff's expert ability to further refine bundles based on factors outside the existing algorithms.

3. **Evaluating project inventory over multiple years adds another dimension of efficiency and value to project bundling.** INDOT has a new research project underway to pull projects from up to 3 years of maintenance inventory in generating the next generation highway contract bundles.

4. **INDOT determined that corridor projects that fell in a 15-25-mile range resulted in the greatest efficiencies.** Smaller groupings did not take full advantage of efficiencies. Longer corridors were more difficult for project engineers to manage and decreased the saving gained from traffic management.

5. **Developing successful bundling applications requires full consideration of formal and informal decision-making rules.** Taking the time to probe and get the unstated decision criteria held by staff in making bundles the table was a crucial element in project success.

6. **More projects will be completed with the same dollars.** Like many states, INDOT's transportation budget is fixed, so the calculated savings are not expected to alter actual budgeted transportation dollars spent. The savings realized will be used to work on more transportation projects with the same budget allocation.

7. **Some decision criteria used by INDOT in the manual process were shown by Machine Learning to be unnecessary constraints.** These decision criteria were adopted to simplify the complexity of the manual bundling process but were not needed and abandoning them allowed the Machine Learning to create even more efficient bundles.

8. **The 15 key data fields are expected to remain consistent and replicable.** While the big savings of early years are likely to diminish over time, it is reasonable to expect this will be offset by algorithms that continue to get smarter and insight that will enable subject matter experts to spot new connections and data relationships.

9. **Adding additional bundling factors can add additional savings**. Examples can include budget considerations, financial considerations, contractor evaluations and other metrics of efficiency will add additional savings that can be generated by the Machine Learning algorithm.

10. **More study is planned.** Based on the results of this process, a parallel research project is being launched to determine if these results can be replicated with INDOT partners in city and county government to jointly create improved transportation corridors and project bundles.

11. **Future enhancements are planned.** These will include development of a statistical model that compares predicted savings to the actual cost savings of bundles. New cost data would be added and refined as it becomes available.

12. **Better projects can improve traffic flow and public safety.** Some of these improvements will be qualitative, others quantitative. Fully verifying savings for transportation projects in Indiana can take up to 5 years to finalize.

Results

- 3,400 projects were analyzed and bundled in minutes vs. days
- More complex and thoroughly vetted bundles were identified with the bundling application
- INDOT staff time to review project bundles was reduced by over 60%
- The bundling application increased savings of 40% over prior method

Comparing identical 2021-2024 reviews, the bundling application identified $107 million more in cost savings than INDOT's prior method

Year	Projects Reviewed	Combined INDOT Savings	Combined Machine Learning Savings	Difference
2021	1155	$85,734,270	$125,041,745	$40,307,476
2022	1051	$80,401,973	$107,359,144	$26,957,171
2023	572	$35,361,682	$58,592,684	$23,232,002
2024	622	$60,872,900	$78,315,713	$17,442,813
TOTAL	**3,400**	**$262,370,825**	**$370,310,288**	**$107,939,462**

BENEFITS

- Better bundles were identified faster
- A broader range of multi-year possibilities were considered
- A standardized approach was used across Districts
- Planning became more inclusive
- Better bundles are expected to improve traffic flow and increase highway safety

Considering Machine Learning for contract bundling? Here are some questions you should ask.

1. How well do you think you are doing at bundling? Will it be painful if it turns out the answer is “not very?” Could uncovering logic errors in existing assumptions be problematic for certain staff or leaders?

2. Does your organization have a data dictionary? Is a starting point of business rules and decision criteria for project bundling available?

3. Are you expecting to hit a button and get a solution? Try using Google translate for an eye-opener on the limitations of this approach. This project demonstrated the importance of detailed and comprehensive front-end work of defining the decision criteria you currently use or would like to use in project bundling. Expect your requirements will require some level of customization. Also, expect to have some misassumptions highlighted by the machine learning.

4. Is your leadership ready for objective criteria and a more inclusive planning process? This approach could pose difficulties for highly centralized organizations or leadership styles that do not value high levels of cooperation and collaboratively.

5. Can the staff and time requirements of your project selection methods be documented? Expectations will change rapidly. Documenting your “as-is” business case before machine learning is important for showing the actual savings.

6. Have you fully considered the informal criteria your agency uses to determine contracting awards? Are there potential “black box” items that need to be included in a bundling algorithm everyone knows, but are difficult to discuss? (i.e., political, contractual, or other contract factors that are not transparent)

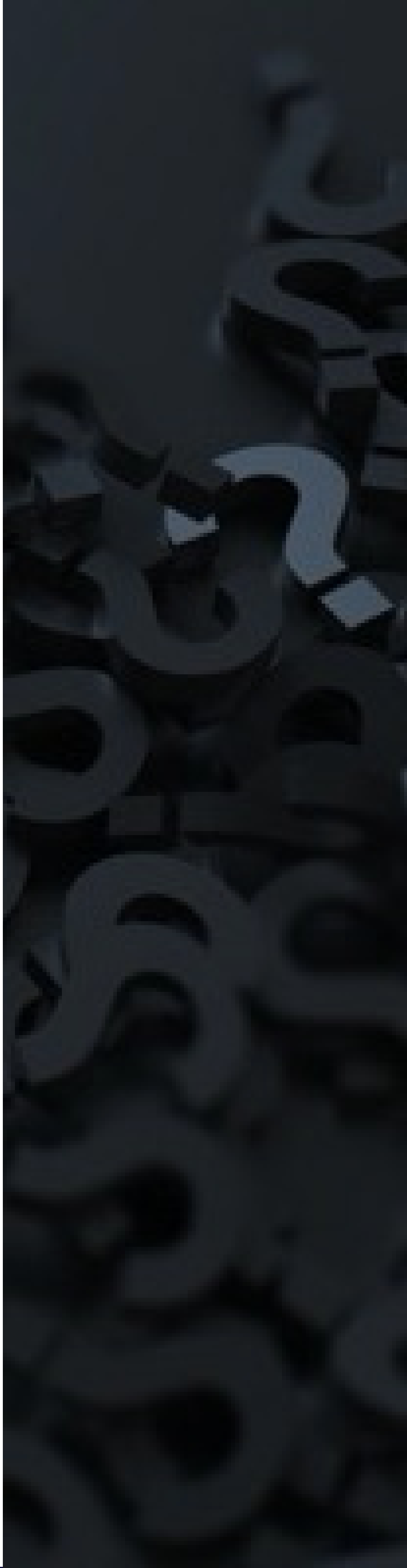
7. Does your asset management database rely on dTIMS? If so, implementation can be straightforward.

8. Are protecting small and specialty businesses and improving minority business participation priorities for your organization? Machine learning can identify opportunities for greater contracting inclusiveness and diversity.

9. What interface requirements will be needed with your organization’s financial system?

10. How will your organization define success: Cost savings? Staff savings? Efficiencies? Other?

11. How will improvements in traffic flow and public safety be quantified?



Brett Boston

A serial entrepreneur, Brett's companies focus on groupware, decision-support and other web-based collaborative solutions. Leveraging these tools, he has produced 2,000+ strategic plans on demanding timelines and worked world-wide with over 800 organizations. Brett is a founder of Foro and Managing partner at Group Solutions.

Vern Herr

An experienced analyst, facilitator analyst and project manager, Vern delivers collaborative solutions to Federal and State construction, recreation and conservation collaborative challenges. Since 1993, he has helped over 200 organizations and their stakeholders identify and address shared challenges. He provided project management and coordination for this research project. Vern is a founder of Foro and a partner at Group Solutions.

Ken Mocabee

Ken has over 23 years' experience building highly interactive applications for a wide range of industries from B2B to consumer-oriented applications. His team's forte is a cloud-native SaaS app that delivers highly effective solutions for business and organizations and brings that experience and approach to the web development disciplines. Ken is a founder of Foro.

Acknowledgments

These results would have been impossible without the leadership of Louis Feagans, Managing Director of Asset Management for Indiana Department of Transportation and close cooperation from Lyndsay Quist, Todd Shields, Mike Wampler, Chris Gentry, Travis Kohl, Doug Burgess and Joel Bump.

The programming, responsiveness and customer experience wizardry of Jeremy Murrish, Mason Mocabee, Peter Brune and Kyle Massey were critical to project success and are gratefully acknowledged.

Box 105603

Atlanta, GA 30348-5603

Tel: 770.757.9828

Email: VHerr@groupsolutions.us

www.teamforo.com