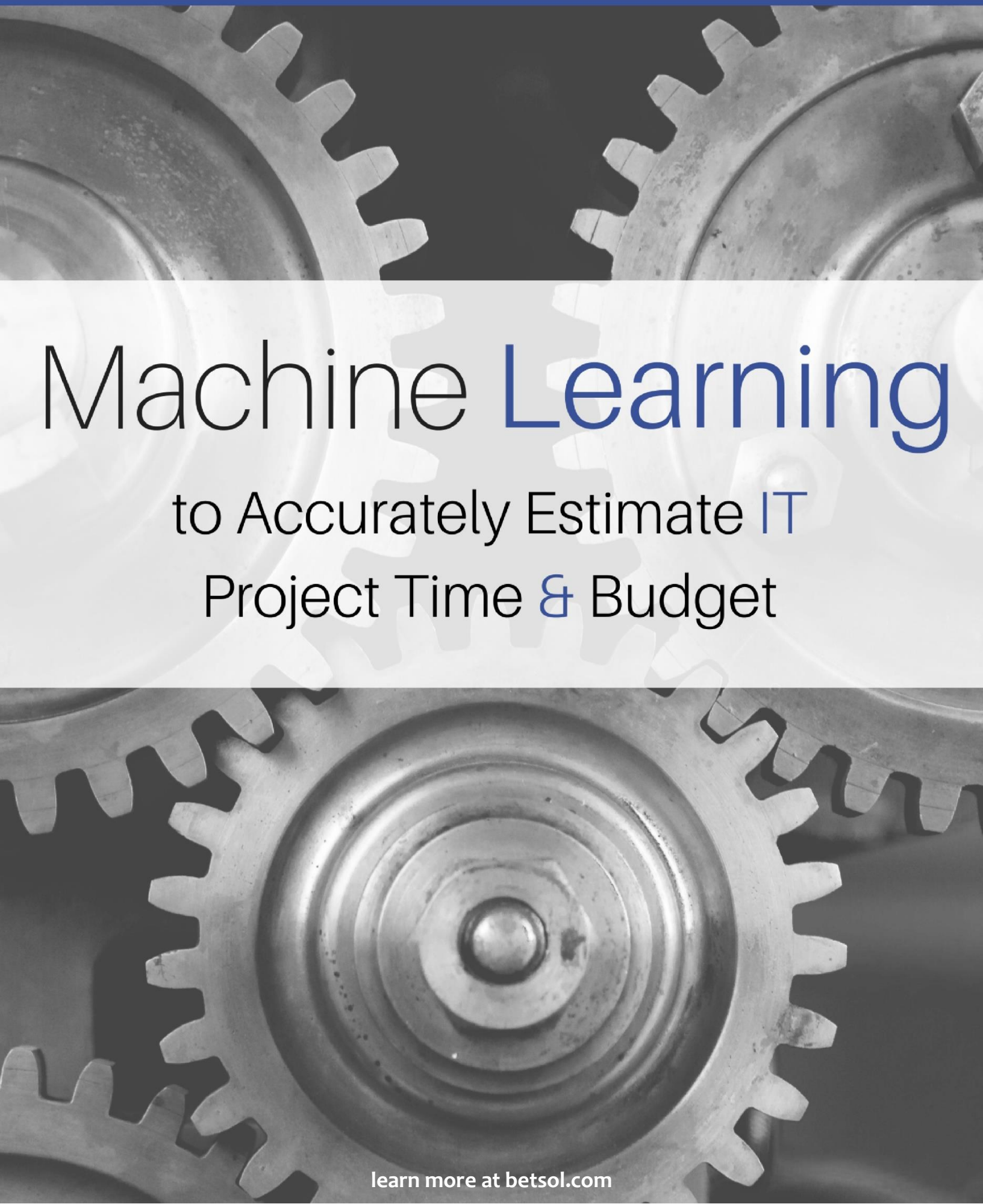
BETSOL®



Machine Learning

to Accurately Estimate IT Project Time & Budget

learn more at betsol.com

Contents

Machine Learning for IT Project Estimation

Is IT Project Estimation a Problem?

What are Common Causes of Poor Estimates?

The Estimate Modeler

Machine Learning to Predict and Improve

Accuracy, Complexity, and Accountability

Tying it all Together: Accurate IT Project Estimates

Terms and Workflow Details

Machine Learning for IT Project Estimation

IT projects notoriously exceed initial time and budget estimates. Accurately estimating these large, complex projects is a challenging task with many unknowns and assumptions. Nevertheless, big-dollar business decisions are made based on IT project estimations, so it is critical to get it right.

Is machine learning the technology that can finally provide the solution?

Absolutely, and we are doing it now.

This white paper will walk you through the concepts involved in implementing machine learning into your standard project processes with a goal of achieving more than a *500% improvement* in time and cost estimation accuracy.

Is IT Project Estimation a Problem?

While you likely have your own experiences with overruns, typical large projects have cost overruns between **560%** and **1,900%** (*Satista*).

The Harvard Business Review performed an in-depth analysis of an SAP system migration project for Levi Strauss. Initial estimates put costs at $5 million, but the project ultimately topped $192.5 million. **That is 3,750% in the wrong direction.**

> "When we broke down the projects' cost overruns, what we found surprised us. The average overrun was 27%—but that figure masks a far more alarming one. Graphing the projects' budget overruns reveals a "fat tail"—a large number of gigantic overages. Fully one in six of the projects we studied was a black swan, with a cost overrun of 200%, on average, and a schedule overrun of almost 70%." *Harvard Business Review*

What are Common Causes of Poor Estimates?

Have you ever sat in a weekly portfolio review meeting, when one of your Program Managers becomes choked up as they tell you that their high-profile project is already 40% over budget and only 50% of the way complete?

If so, have you ever asked yourself, "*How did we get so far away from the original estimate?*"

When we ask IT Project and Program Managers what they lack when creating an estimate, they all express similar struggles.

1) **Poor Information:** Building estimates requires involving numerous people throughout the business. When Program Managers seek cross-department involvement, business units are not typically held accountable for the accuracy of their inputs. The Program Manager is provided with assumptions; often hastily-derived assumptions.
2) **Poor tools:** We have also found that most estimates are still being created in Excel, oftentimes, by a single "Excel Ninja" person. Single calculation errors can have enormous ripple effects. Excel Ninjas introduce major bottlenecks and risks. Despite these risks, Excel continues to be used because it is so easily customized and can flex as the business shifts.

The Estimate Modeler

An Estimate Modeler solves both of these common problems. It creates accurate base estimates on the fly while interacting with experts to refine numbers. It holds organizations accountable and learns from mistakes. It also provides flexible, modern tools to better analyze and present data with lower risk of error.

- **Analytical Dashboards** to help avoid surprises
- **Cloud-Based** solutions for ongoing collaboration
- **Centralized Project Request** processes for continuity

Machine Learning to Predict and Improve

Estimate Modelers generate large amounts of data, which makes it a ripe field to apply Machine Learning. The goal is to help estimators make smarter decisions based on historical and actual data, guiding them to better estimates.

The solution centralizes the estimation process allowing Estimators to collaborate with Subject Matter Experts (SME) to refine estimates. The estimate is built up from a default template that is enhanced by user-controlled parameters.

And because we realize information improves over time, the solution uses *complexity levels* to refine estimates throughout the planning process.

Accuracy, Complexity, and Accountability

Following are the key components of the Estimate Modeler solution.

Collaboration

The right Estimate Modeler solution supports collaboration without having to use other tools. Internal triggers send information requests to SMEs. The SMEs access the solution and provide their input on the information requested. The solution embraces accountability, giving estimators the ability to assign due dates, track progress, and send reminders.

User-controlled parameters

Default parameters are used to initially gather information about projects. Based on big data, we were able to create a list of over 100 parameters to get the model dialed in. These parameters are interchangeable across different industries and business sizes, allowing estimators and SMEs to rapidly begin the process.

Refining estimates with complexity

Estimates are fluid and should be refined as more is learned about the project. We address this with "complexity levels." Complexity levels and estimate parameters have a direct correlation.

As you go up in complexity level, you go deeper into the more granular parameters. The complexity levels also create benchmarks for the self-learning model to process against.

Machine learning

As this Estimate Modeler runs, Machine Learning processes data and suggests new defaults to the estimator. Linear Regression processes and compares different data sets including: current estimate, historical estimate, and actual hours billed. Bayesian techniques fill in gaps and suggest new variables.

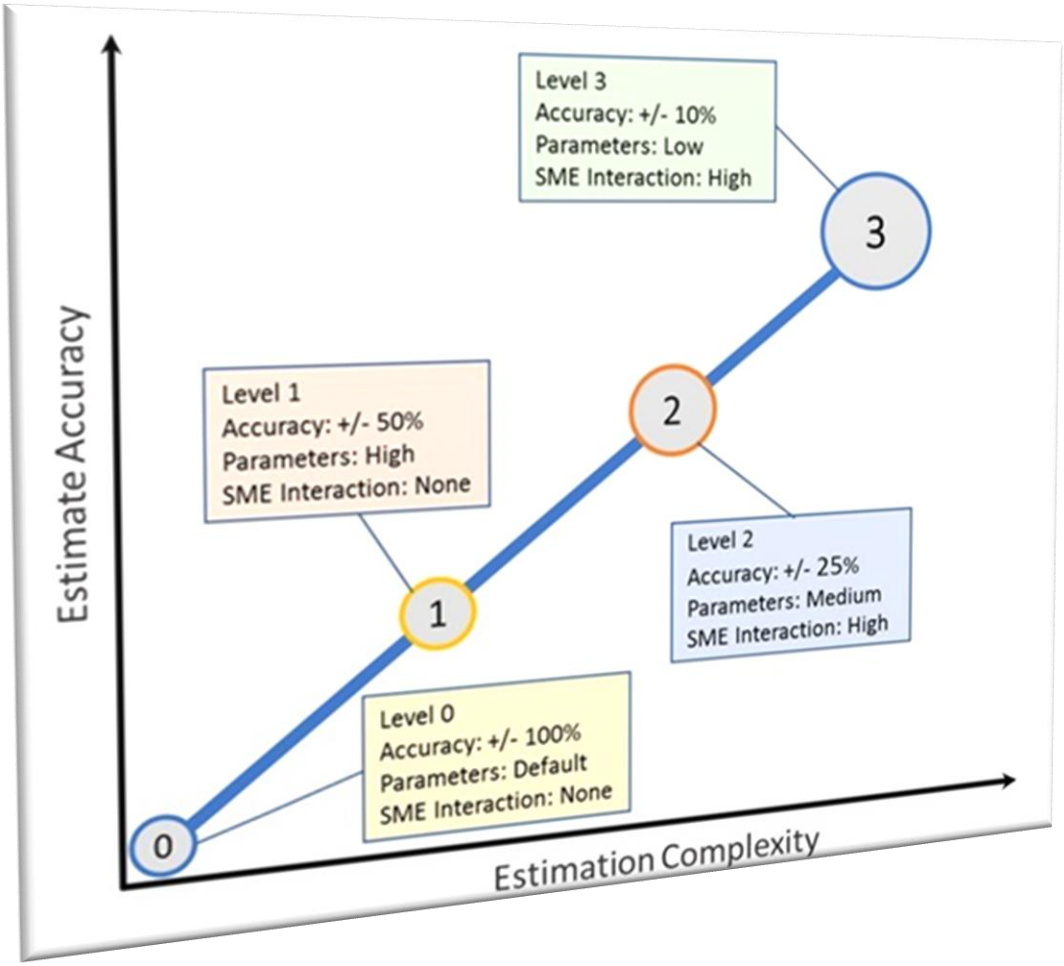
The Machine Learning Estimate Modeler is built to be suggestive–the estimator still has the final say and ability to override inputs.

Project and Program Reporting

Dashboards allow managers to quickly view project health and financial reports.

User Management and Role Based Access Control

Users interact with only the components that apply to them. When up and running, this solution is the framework to help Enterprise Estimators be massively more accurate–whether the project is **$100K or $100M.**



Tying it all Together: Accurate IT Project Estimates

In order to bring the solution to life, we have utilized open-source tools and industry accepted methodologies. We have built libraries, functions, and a modern user interface on top of the solution in order to remove the need for a developer once it is stood up.

Terms and Workflow Details

Parameters – information about tasks, data, or other assets that effect a projects timeline.

- High – A high-level parameter should not require additional information to be gathered from Subject Matter Experts (SME). These parameters are open-ended questions: Who is the client; What is the skill level of the PM/s; Is the project scope defined.
- Medium level – A parameter that would require input from an SME, often times requiring additional research by the SME. These parameters are still open-ended questions at this point: Skill Level of PM with Tools X, Y, Z; what the data model is; what the data source is.
- Low level – A parameter that would require input from an SME. These parameters will require research and some level of analysis. These parameters ask for quantities: # of Reports; # of Cubes; # of Data Marts.

Complexity Levels– The higher level of complexity = higher level of accuracy.

- Level 0 – This will be the default estimate that is created at the time a project is requested. The estimator will have little information about project specifics so all hours are set at a department level. The project is still in the Idea Phase. This estimate can be +/- 100% from actuals.
- Level 1 – The Estimator has started gathering more information about the project and it has moved to become an Approved Idea. The estimator will create a new level of the estimate answering questions about high level parameters. This estimate can be +/- 50% from actuals.
- Level 2 – At this point the estimator will have to reach out to SMEs to answer questions about the project complexity. The project has entered the definition stage at this point. This estimate can be +/- 25% from actuals.
- Level 3 – The estimator will continue to interact with SMEs in order to refine project information and deliverables. This will be the last estimate that is made before the project launches into development phase. This estimate can be +/- 10% from actual.