DECISION
MANAGEMENT
SOLUTIONS

By James Taylor

CONTENTS

CRISP-DM & Decision Management

Creating a decision-centric, repeatable approach

Smart companies know they need a methodology to succeed with new technologies. Advanced analytics and data science projects are no exception: simply investing in big data infrastructure and buying new tools won't suddenly make your organization's decisions smarter.

CRISP-DM: An Overview

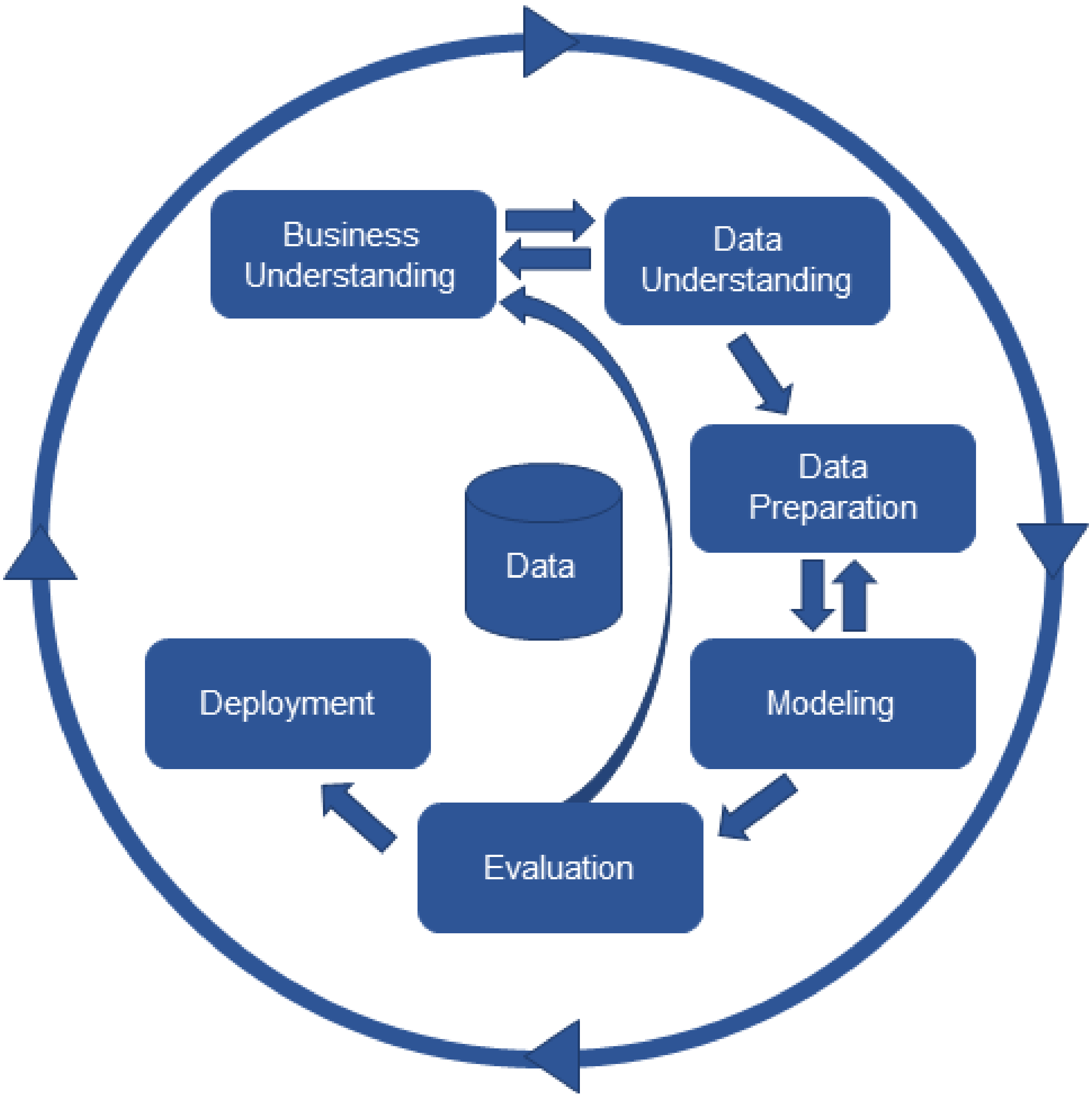
CRISP-DM - the **CR**oss **I**ndustry **S**tandard **P**rocess for **D**ata **M**ining is a great framework for advanced analytics and data science (data mining) projects, especially for organizations that lack experience with these kinds of projects and don't yet know what kind of analytics will help with which decisions. Conceived in late 1996 and developed by a consortium formed by Daimler-Benz, SPSS and NCR in 1997. CRISP-DM 1.0 was released in 2000. Despite being conceived over 20 years ago, it is still the most popular and effective methodology for advanced analytics

CRISP-DM begins by establishing the business problem and understanding the available data. The data is then prepared and analytically modeled iteratively until a result is achieved that can be evaluated and deployed. The whole approach is designed to be both iterative and repeatable.

The six phases are:

- Business Understanding
  Determine business objectives, assess the situation, determine your analytic goals and produce a project plan.
- Data Understanding
  Collect, describe, explore and verify the quality of the data available.
- Data Preparation
  Select, clean, construct, integrate and format the data needed.

Figure 1. The major phases of CRISP-DM



- Modeling
  Select modeling technique(s), generate a test design, build and assess the model
- Evaluation
  Evaluate results and process to determine next steps.
- Deployment
  Plan deployment and monitoring/maintenance approach, finalize and deploy.

All of this is then iterated as necessary to continuously improve the analytic results being delivered.

There are three key aspects of CRISP-DM where Decision Management and decision modeling play a critical role:

- Business Understanding
- IT Engagement
- Deploying and Delivering Business Value

Business Understanding

One of the most important elements of CRISP-DM—and perhaps one that is too often neglected by data scientists keen to get started with the data—is the focus on having business understanding first, before any analytical work begins. The core idea is that the business objective and situation must drive the project rather than the data science. The business understanding phase is supposed to generate the following work products:

- Background information
- Glossary of terminology
- Risks and contingencies
- Business objectives and success criteria
- Requirements, assumptions, and constraints
- Inventory of resources
- Cost-benefit analysis
- Data-mining success criteria and goals

These are all great deliverables. The question for the project team using CRISP-DM is how to capture the requirements, assumptions, constraints, and available resources to make the data-mining success criteria clear and to match those criteria to the true business objective.

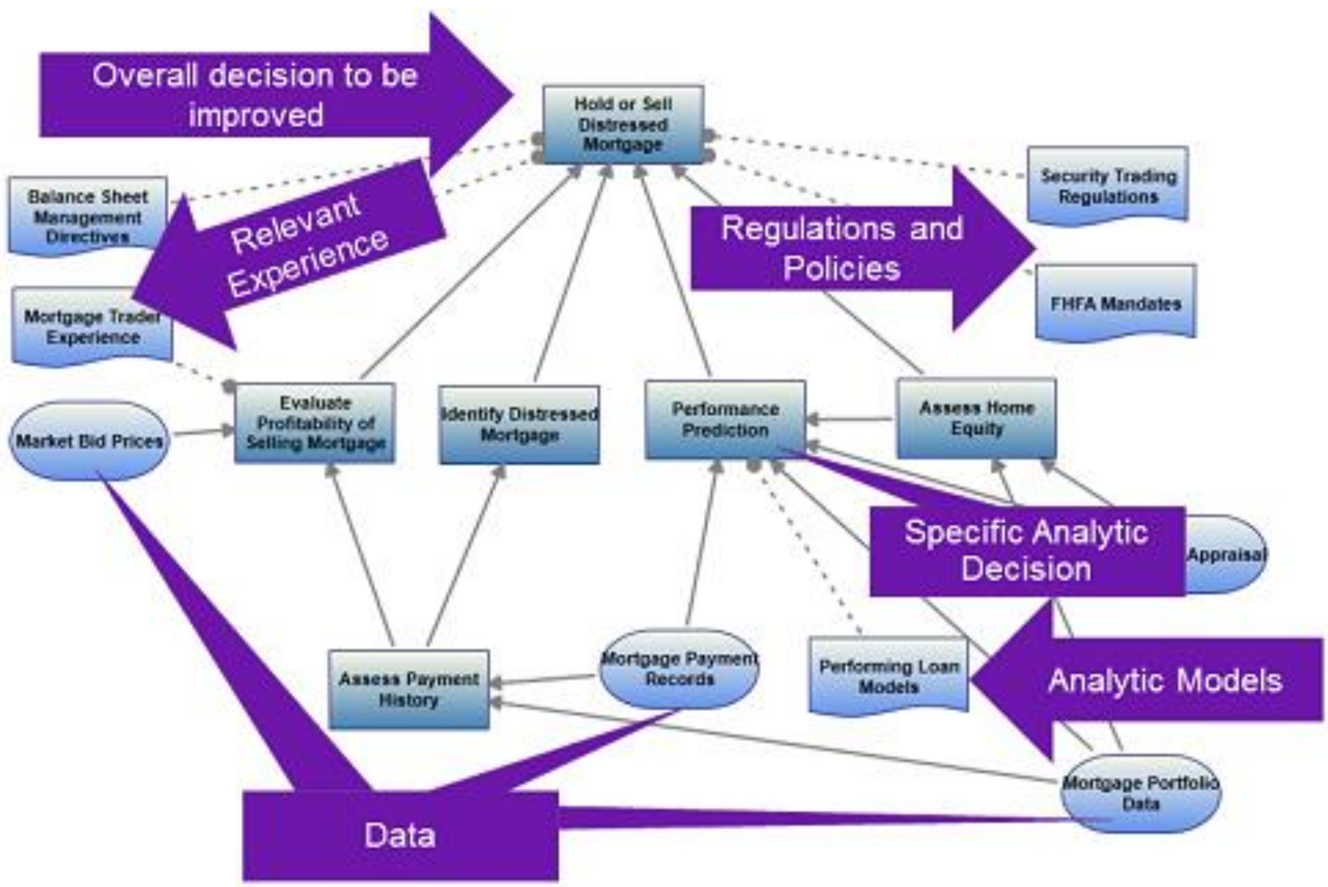
Although it is possible to write a document that does this, it is much more effective to build a model to do so. What kind of model? Advanced analytics projects should be focused on improving decision making.

> *"If analytics does not lead to more informed decisions and more effective actions, then why do it at all?"*
>
> *— Mike Gualtieri, Forrester Research*

In recent years, decision modeling has become an increasingly mainstream technique. There is a published standard for decision models—Decision Model and Notation (DMN)—as well as several books on the subject. DMN decision models, such as that shown in Figure 2, take a targeted decision and break it down into sub-decisions (and sub-sub-decisions, etc.). Each decision is described in terms of the specific question that must be answered and is linked to its previous decisions as well as to the required input data. This defines the information needed to make the decision—that is, either data input to the decision or information resulting from making other decisions. Available knowledge sources (e.g., policies, regulations, best practices) are also identified and linked to decisions to make it clear how those decisions should be made. The whole model is linked to the business objectives the decisions impact, the organizations involved in making the decisions, and the processes during which the decisions must be made.

Critically for CRISP-DM projects, decision models can also include analytics or data science outputs as knowledge sources so that the role of the project's data science outputs (e.g., regression models, neural networks, decision trees) in decision making can be clearly shown.

Figure 2. A DMN Decision Model Showing the Specific Role of an Analytic Model



Project teams that build decision models in the business understanding phase bring clarity to the problem their data science is intended to solve. Decision models keep the focus squarely on business decision making, not on the technologies involved, as well as clearly show what the analytics should help decision makers do. Decision models also connect the proposed analytics to business objectives and business processes, so it is evident when the analytics will be used and whether it has had a positive business impact. Most important, decision models offer benefits throughout the project.

IT Engagement

Project teams that build a decision model to define their requirements and capture business understanding bring clarity to the problem they are solving. Moreover, a decision model can ensure ongoing and effective engagement of both business and IT professionals in the analytical core of a project.

Analytical model development is highly iterative—data scientists working on the model constantly evaluate the available data in new ways, try different sampling algorithms, apply different analytical techniques, and combine data differently. This is essential to the effective development of an analytical model and the tasks are quite technical, with little or no role for the business or IT practitioners on the team. The tasks also make up the bulk of the project effort, covering the data understanding, data preparation, and modeling stages in CRISP-DM.

This creates a singular danger—that the analytics members of the team will drift away from the business and IT members. Because these tasks are not broadly understood or accessible, the business and IT partners cannot participate. Each iteration can result in the analytics team being further disconnected until it gets completely out of sync. A decision model helps prevent this by giving the analytics team a framework within which each iteration must fit—if the iteration is not moving toward a better decision, then it is not helping.

A decision model also makes it clear when the analytics team needs to reengage the business. Sometimes the analytics team cannot find a way to build the analytical model originally envisioned. In the process, the team may well find there is something else it can predict or describe. At this point, it can bring up the decision model, see what changes would be required to take advantage of the new analytics, and engage with business and IT partners to see if such a change is practical. This enables the business perspective to be iterated in parallel with the analytics perspective.

Projects can also suffer from what can be described as the “shiny object” problem. There are many exciting analytics technologies available with more coming online all the time. Cognitive and artificial intelligence techniques allow analytics to be developed against documents, email, and even images. Machine-learning and deep-learning techniques create analytical insight from noisier and more complex data sets. In addition, many new data sources are available—both internally, as organizations adopt data lakes, and externally, as the application programming interface (API) economy makes more data available for purchase.

All these new analytics opportunities can be distracting— they are bright shiny objects that can tempt the project team from its path. This is true of both business and analytics professionals involved in the project. Analytics and data science professionals are deluged with articles about new algorithms and new data sources from their peers and technical communities. They naturally want to use the coolest, newest approaches. Business people, too, see articles aimed at lay readers detailing the business value some other organization realized from a certain kind of data or algorithm. Naturally, they seek the same value from their project.

A decision model provides a set of guide rails and signposts for the project. It details which decision is to be improved, what a better decision looks like, who makes the decision, and where analytics or data science is expected to make a difference in making the decision. This provides an effective touchstone for the project team for any new capability: Given our focus on this decision, will this new technology or data help us achieve our business objectives? If no, it can be left to one side (at least for now). If yes, the decision model gives clear guidance and justification for its adoption.

Deploying and Delivering Business Value

One of the key advantages of the CRISP-DM approach is that it includes evaluation and deployment. Organizations that succeed with advanced analytics know that being able to operationalize their analytics is critical to success. Only by embedding their analytics into operational workflows and taking action based on those analytics can you create value.

An analytic team applying CRISP-DM begun their project by documenting their business understanding – ideally as a decision model based on the Decision Model and Notation standard (DMN) such as that illustrated in Figure 2. This captures the requirements, assumptions and constraints of the decision to be improved. It also ties the project's *analytical* success criteria to the true *business* success criteria, ensuring the team knows how good their analytic model must be if it is to add business value.

Analytical model development is highly iterative. Multiple modeling approaches are tried, refined, rejected and combined to come up with an effective model. There is a danger, however, that the analytic team can become over-focused on their analytic accuracy and lose sight of the business objective.

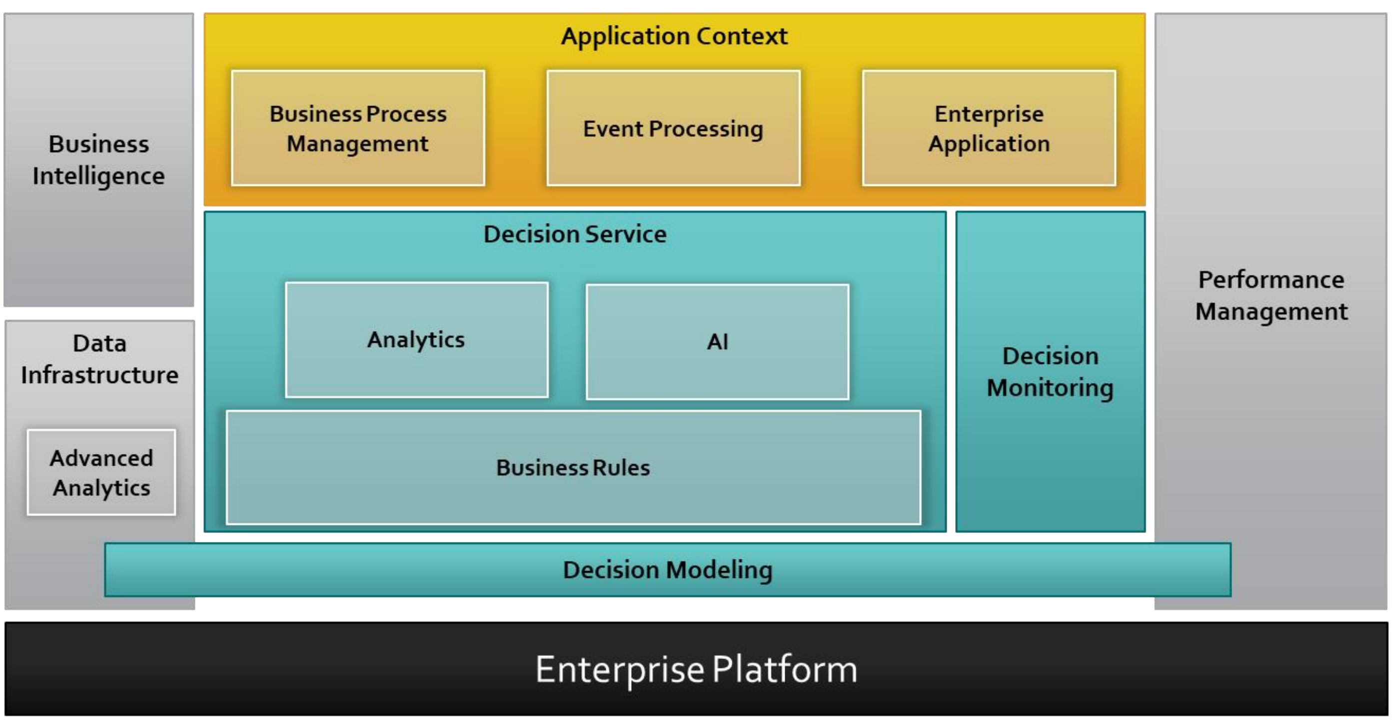
Analytic teams sometimes see that they could make a model more accurate – turning a yes/no answer into quintiles or deciles for instance – and start developing an analytic model that is as accurate as the data allows. While this seems harmless – after all a more accurate model is surely better – it takes time and resources that might be better spent elsewhere.

Sometimes, however, the data does not support a very accurate model. At this point the analytic team may decide to abandon the effort because only a poor-quality model seems likely to be developed or they may continue to work hard extracting whatever accuracy they can, even though there is a clear sense that there is an upper limit to the accuracy that can be developed.

A decision model provides a touchstone at this point. The analytic team can refer to the decision model to see what accuracy of model is needed to actually improve decision-making. They can see that the model they have developed is already accurate enough to discriminate between the options outlined in the decision model and so stop working on the model. They can see that a model is unlikely to hit the level needed to be used in the decision and so return to the business understanding step, working with their business partners to see how decision-making might be changed to reflect the analytic models that are possible. The decision model ensures that the data scientists working on the analytic model can conduct a *business* evaluation, not just an analytic one.

Some analytic teams think they are done once they have a model that meets this business evaluation. Yet the business is still not getting any value from their work. Until and unless the analytic is deployed, embedded in operational workflows and resulting in different actions being taken, no value has been added. Successful analytic teams continue to work with their business and IT partners to ensure that deployment – the final step in CRISP-DM – is successful.

Figure 3. Decision Management Technology Architecture



Deployment generally involves working with IT to embed the analytic model, the algorithm developed, into a production environment. As shown in Figure 3, this involves coordinating business rules, analytics and AI components into a decision service that supports various application contexts.

Once the model is deployed, the final step is to work with the business and operations teams to ensure that change happens, that behavior changes to exploit the new analytic approach.

The decision model created at the beginning of the project ensures effective deployment. A decision model shows what the balance is between the analytic model and explicit decision logic or business rules resulting from policies or regulations that must be enforced when the decision is made. Because decision models are also a great requirements technique for business rules, a single model shows how business rules and analytics are combined to make a data-driven, compliant decision. The decision model also shows exactly what data is used where in the decision, easing the necessary data integration for the analytic.

The decision model also acts as a training and implementation framework for those involved. With decisions mapped to different organizations and roles as well as a clear link to business objectives, a decision model shows the organization how it needs to make the new analytical decision. The visual nature of a decision model makes it easy to understand and decision models are very effective training tools. The balance between automated and manual decisions as well as that between judgmental and analytic decisions is shown clearly. The decision model provides a roadmap for the necessary organizational change.

Finally, the decision model ties the organization's data to its decision-making and to the performance management environment. Decision monitoring is based on tracking how the decision was made (in terms of the decision model) and then linking this decision-making data to the business outcomes being tracked in the performance management environment.

Conclusion

Adopting advanced analytics can seem daunting and technology driven. Adopting CRISP-DM and decision modeling allows teams new to these technologies to make a strong business case for them and tie proposed projects to real business value.

- Decision modeling is an accessible technique that enables team members with business domain expertise but not data science skills to be active participants.
- Keeping a decision model front and center and iterating it when necessary allows data science projects to keep business and IT practitioners engaged and relevant.
- A decision models ensures the right analytic gets built, gives the analytic team a business goal to evaluate their progress against, and ensures that the resulting analytic model can be deployed and effectively used.

Decision models and the CRISP-DM methodology help you ensure you get business value from your analytic investments.

Learn More

Decision Management Solutions has extensive experience helping organizations learn, adopt and implement Decision Management. Our clients are leading companies in insurance, banking, manufacturing, telecommunications, travel and leisure, health management, and retail.

- Client Case Studies
- Our Services
- Contact Us Today for a Free Consultation

CONTACT US

Decision Management Solutions specializes in helping organizations build decision-centric, action-oriented systems and processes using decision management, business rules and advanced analytic technologies.

www.decisionmanagementsolutions.com Email : info@decisionmanagementsolutions.com