

CRISP-DM stage five – evaluation

Evaluate results

Task

Previous evaluation steps dealt with factors such as the accuracy and generality of the model. This step assesses the degree to which the model meets the business objectives and seeks to determine if there is some business reason why this model is deficient. Another option is to test the model(s) on test applications in the real application, if time and budget constraints permit. Moreover, evaluation also assesses other data mining results generated. Data mining results involve models that are necessarily related to the original business objectives and all other findings that are not necessarily related to the original business objectives, but might also unveil additional challenges, information, or hints for future directions.

Output

- **Assessment of data mining results** – summarise assessment results in terms of business success criteria, including a final statement regarding whether the project already meets the initial business objectives.
- **Approved models** – after assessing models with respect to business success criteria, the generated models that meet the selected criteria become the approved models.

Review process

Task

At this point, the resulting models appear to be satisfactory and to satisfy business needs. It is now appropriate to do a more thorough review of the data mining engagement in order to determine if there is any important factor or task that has somehow been overlooked. This review also covers quality assurance issues—for example: did we correctly build the model? Did we use only the attributes that we can use and that are available for future analyses?

Output

- **Review of process** – summarise the process review and highlight activities that have been missed and those that should be repeated.

Determine next steps

Task

Depending on the results of the assessment and the process review, the project team decides how to proceed. The team decides whether to finish this project and move on to deployment, initiate further iterations, or set up new data mining projects. This task includes analyses of remaining resources and budget, which may influence the decisions.

Output

- **List of possible actions** – list the potential further actions, along with the reasons for and against each option.
- **Decision** – describe the decision as to how to proceed, along with the rationale.