

Cross-Industry

Data standardization in Ethiopia

Standard practices for data mining six stages in CRISP=DM methodology

1. Business understanding

Business requirement documents or functional requirement specifications. It consists of document version control, the stakeholders which are involved in this particular use cases, what is the confidentiality statement, what is the business requirement, what are the business objectives, what are the various model that can be built to solve this particular business [rpblm, what are the data elements which are required to solve this particular problem and how these data elements can be sourced from various data sources.

2. Data understanding

Data elements are mapped to all those source systems from where we can get these data.

3. Data preparation

All the mappings are done. The business departments are consulted to tell what are the various models that can be used for solving the particular case in use. For example in analytics there can be many models which can be used where models like random forecast precision trees, care nearest neighbors support machine,

4. Data modeling

Various iteration can be shown and there is a precision and accuracy level that people measure to come up with the model out puts. These model out puts are evaluated I the fifth stage of evaluation and finally deployment is done.

5. Evaluation and

6. Deployment

THE SAMPLE ANALYTICAL RESEARCH

1. Churn prediction

2. Fraud control

3. Customer segmentation and profiling for the right mix of the products for the customers.

4. Loan default prediction

5. NPA segmentation and profiling

6. Product profitability and product pricing