

2.2.3 Data Analytics Lifecycle

The lifecycle consists of several phases

1. Discovery

- Learn about the problem domain, conduct interviews, study sources of data.
- Study the past problems that were solved.
- Formulate initial hypothesis, ideas that can be tested.

2. Data preparation

- Execute ELT (extract, load, transform) to get data into the data warehouse.
- Conditioning through cleaning and normalization.
- Conversion from categorical to numerical or one-hot vector.

3. Model planning

- Selection of methods, techniques and processing pipelines.
- Ensuring proper structure of datasets.
- Exploring data to select important variables.
- Selecting models appropriate for the nature of data available (structured or unstructured).

4. Model building

- Develop testing and training datasets.
- Implement the model.
- Ensure availability of hardware requirements.
- Execute the model.

5. Communication of results

- Evaluate business value of the results.
- List the key findings and summaries.
- Suggest future improvements.
- Prepare documentation.

6. Operationalize

- Start using the model in the production environment.
- Deploy the model in controlled way.
- Provide concise responses for effective real-time decision making.