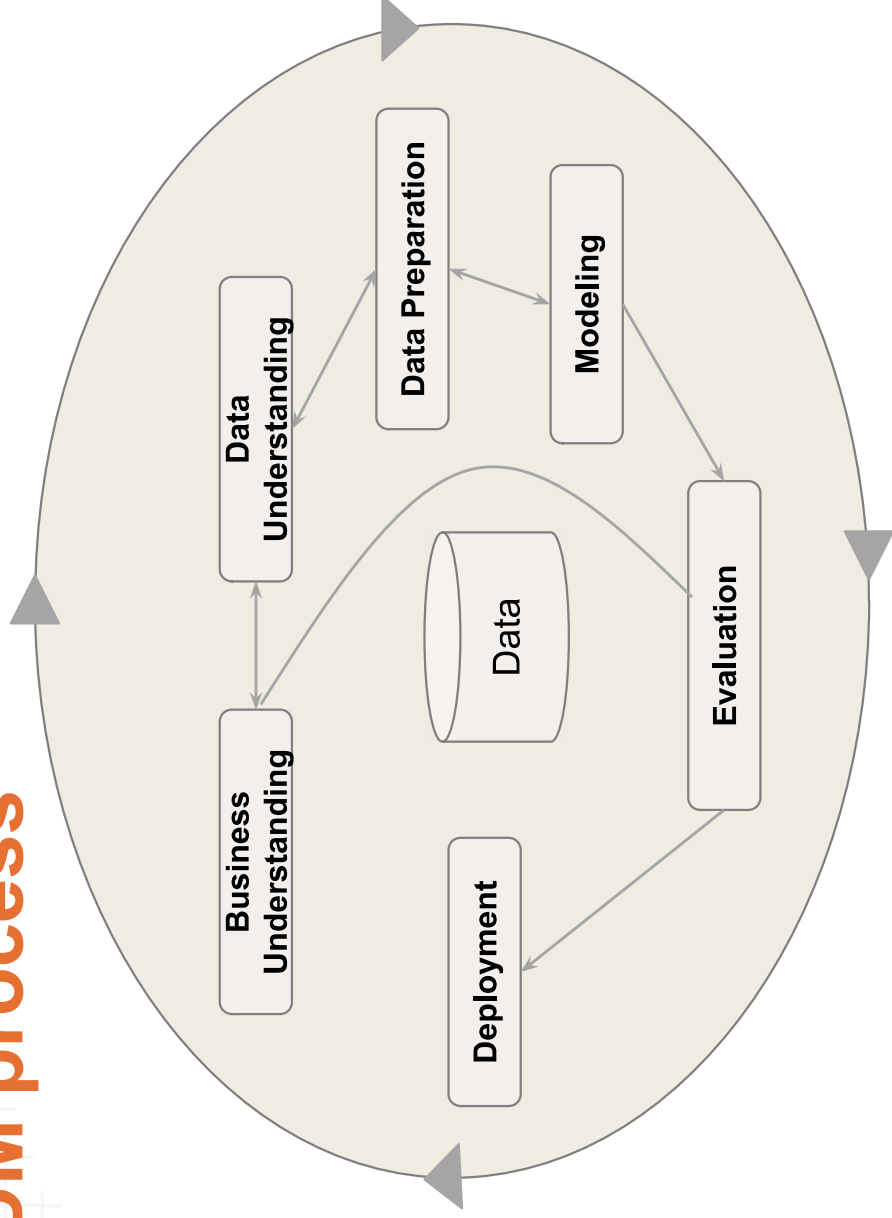


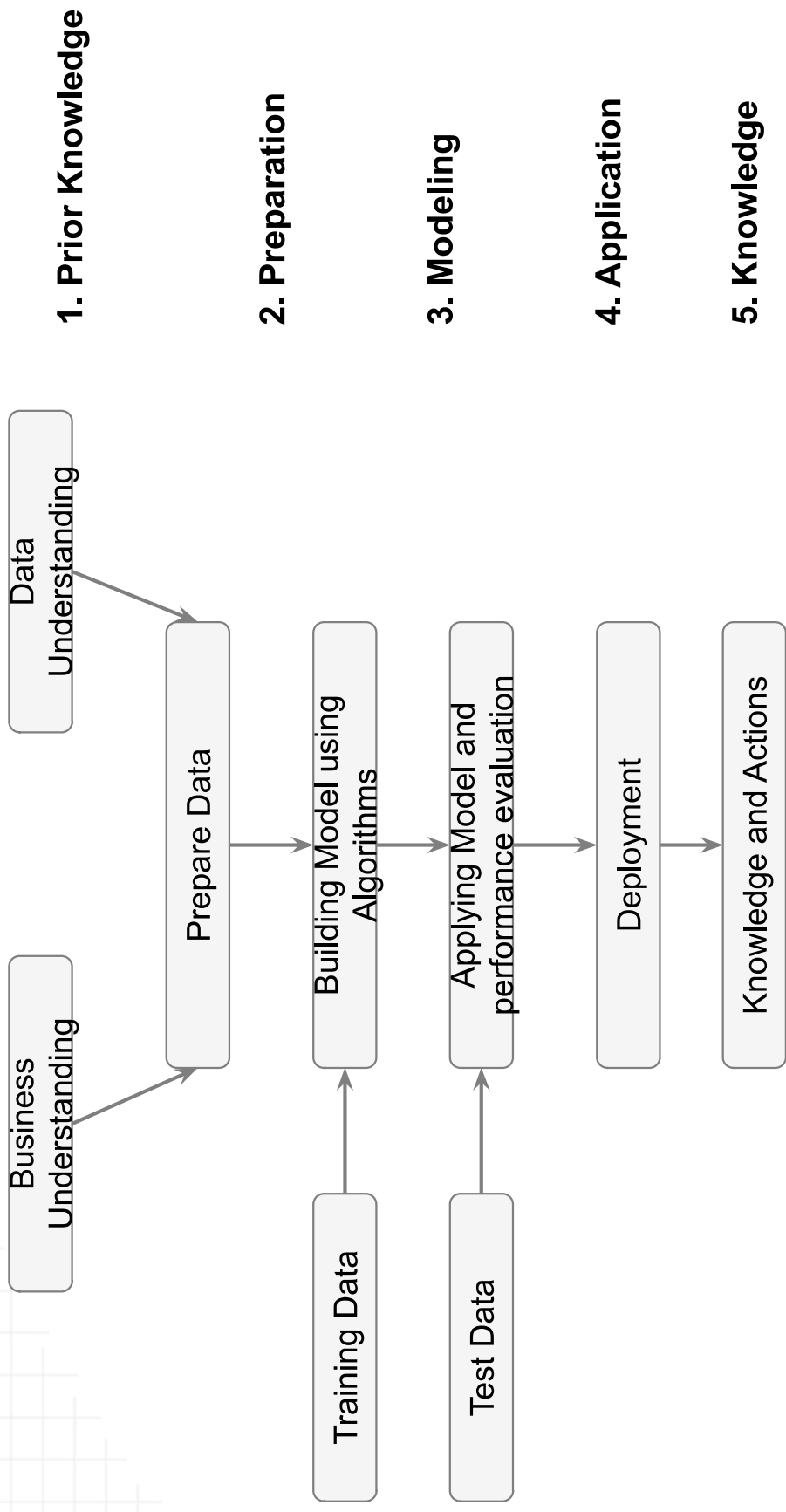
The background is a grayscale abstract image. It features a grid of thin, light gray lines that intersect to form a pattern of small squares. Overlaid on this grid are various numbers in different sizes and orientations. Some numbers are bold and clear, while others are faint and blurry. The numbers include digits from 0 to 9, as well as some letters like 'a' and 'b'. The overall effect is one of data and complexity. A dark gray horizontal bar is positioned across the middle of the image, containing the text '2. Data Mining Process' in white. The bar has a small orange square at its left end.

2. Data Mining Process

CRISP DM process



Process



1. Prior Knowledge

Gaining information on:

- Objective of the problem
- Subject area of the problem
- Data

Table 2.1 Data Set		
Borrower ID	Credit Score	Interest Rate
01	500	7.31%
02	600	6.70%
03	700	5.95%
04	700	6.40%
05	800	5.40%
06	800	5.70%
07	750	5.90%
08	550	7.00%
09	650	6.50%
10	825	5.70%

2. Data Preparation

Data Exploration

Data quality

Handling missing values

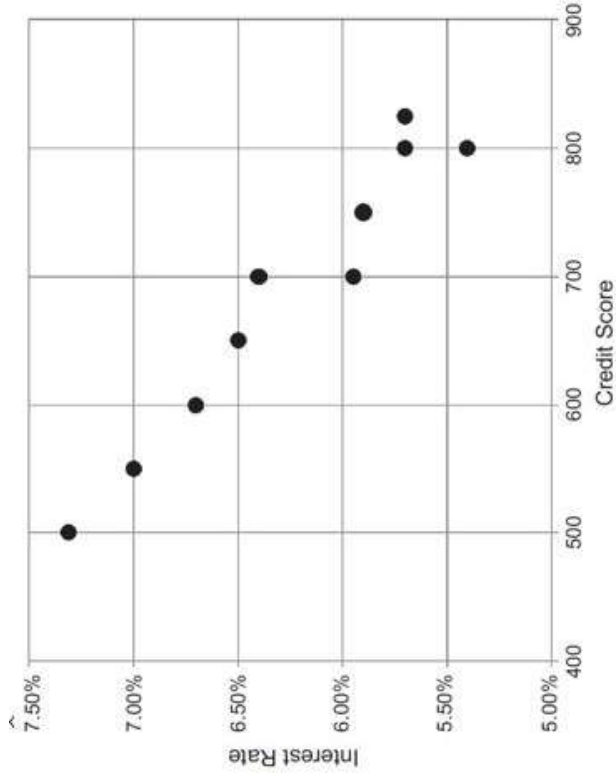
Data type conversion

Transformation

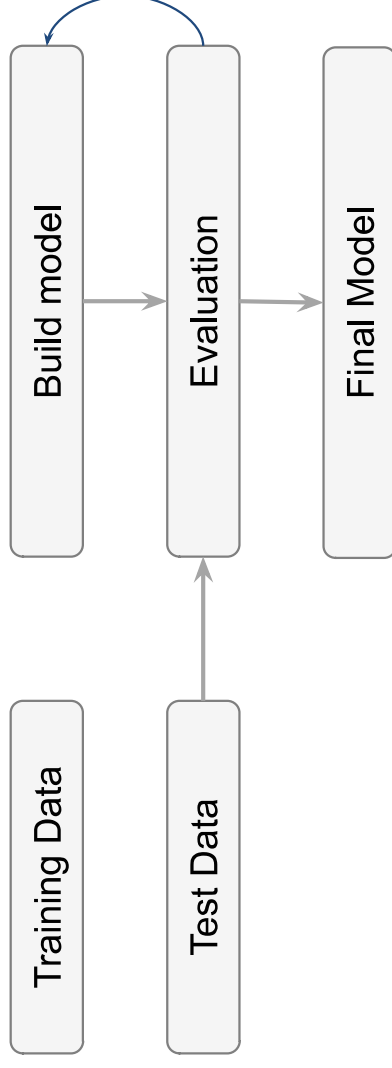
Outliers

Feature selection

Sampling



3. Modeling



3. Modeling

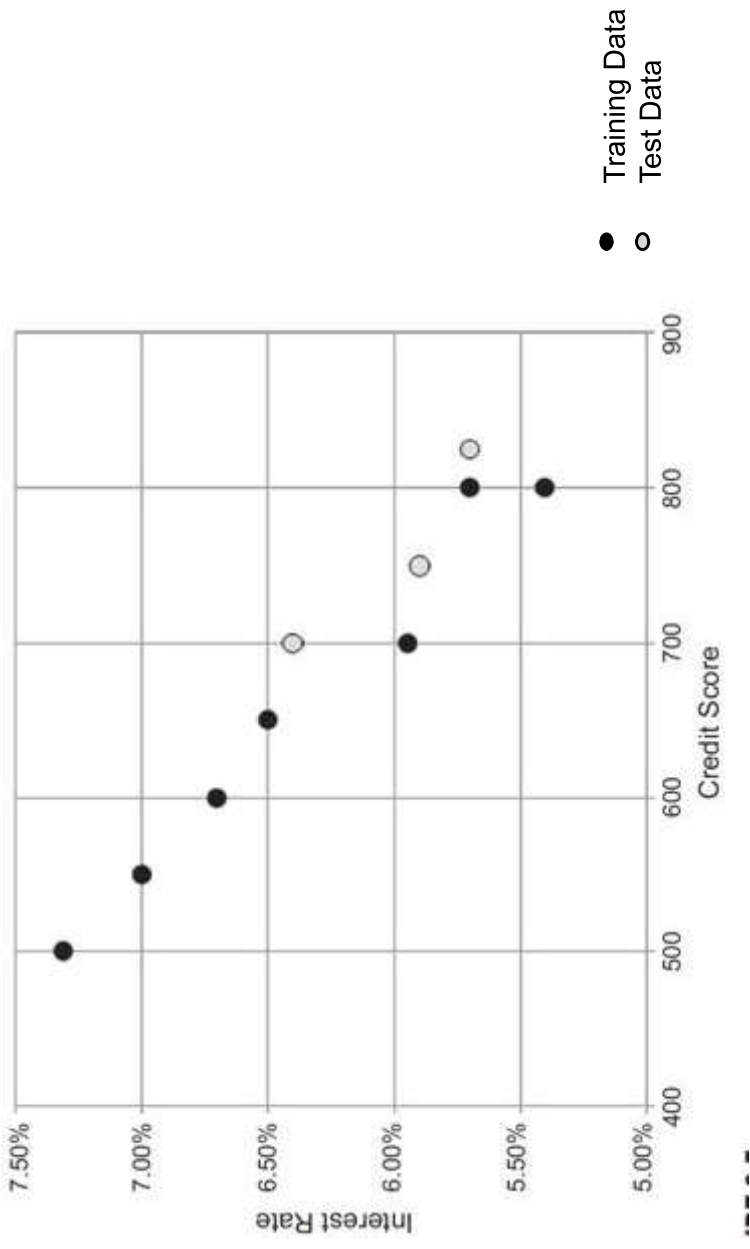
Splitting training and test data sets

Table 2.3 Training Data Set			
Borrower	Credit Score (X)	Interest Rate (Y)	
01	500	7.31 %	
02	600	6.70%	
03	700	5.95%	
05	800	5.40%	
06	800	5.70%	
08	550	7.00%	
09	650	6.50%	

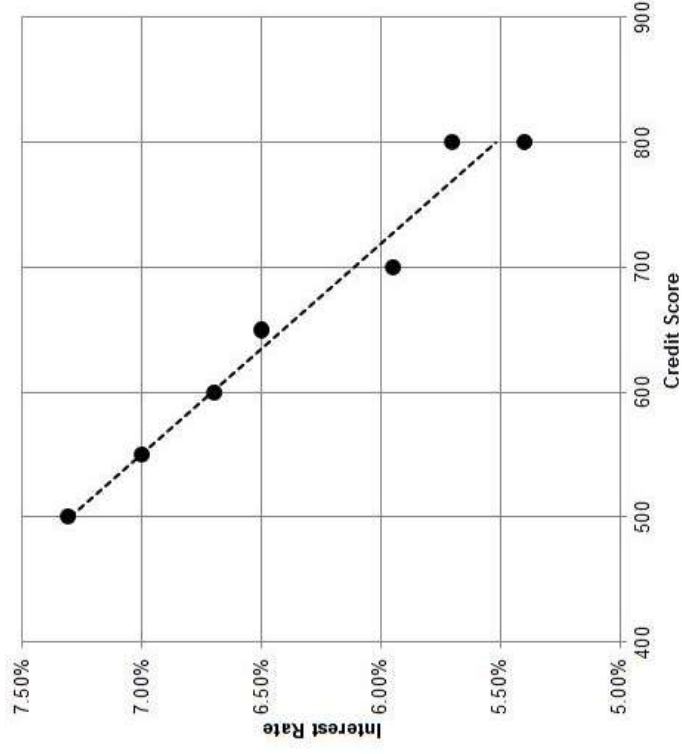
Table 2.4 Test Data Set		
Borrower	Credit Score (X)	Interest Rate (Y)
04	700	6.40%
07	750	5.90%
10	825	5.70%

3. Modeling

Splitting training and test data sets



3. Modeling



$$y = 0.1 + \frac{6}{100,000}x$$

3. Modeling

Evaluation of test dataset

Table 2.5 Evaluation of Test Data Set				
Borrower	Credit Score (X)	Interest Rate (Y)	Model Predicted (Y)	Model Error
04	700	6.40%	6.11%	-0.29%
07	750	5.90%	5.81%	-0.09%
10	825	5.70%	5.37%	-0.33%

3. Application

Product readiness
Technical integration
Model response time
Remodeling
Assimilation