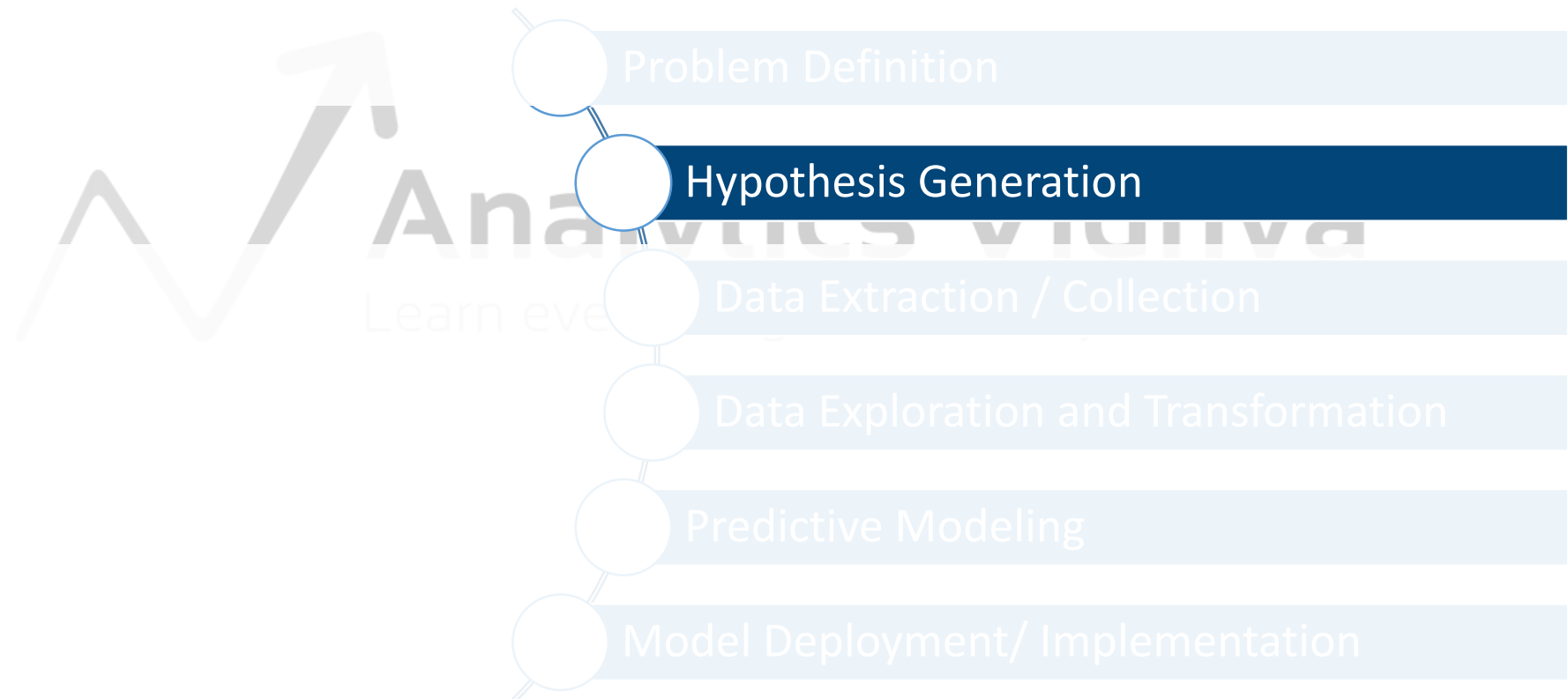




Predictive Modeling

Stages of Predictive Modeling

We can broadly divide the model building life cycle in six stages:



Hypothesis generation

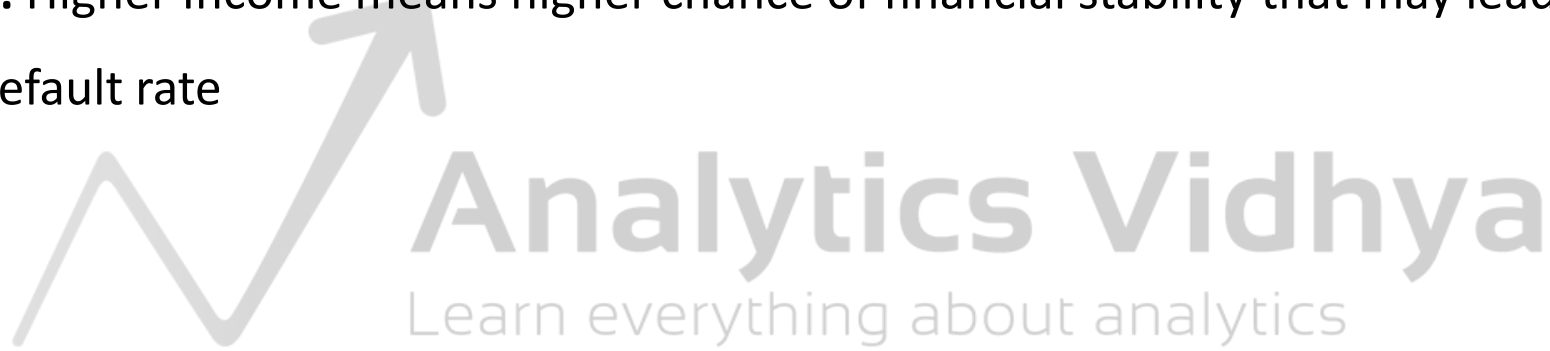
List down all possible variables, which might influence problem objective

Analytics Vidhya
Learn everything about analytics

Hypothesis generation

Example: What are factors, those can impact the customer will default or not?

- **Income:** Higher Income means higher chance of financial stability that may lead to lower default rate



Hypothesis generation

Example: What are factors, those can impact the customer will default or not?

- **Income:** Higher Income means higher chance of financial stability that may lead to lower default rate
- **Job Type:** If job is more stable that can also lead to lower default rate. Example: Front End Sales job is less stable than data science job

Hypothesis generation

Example: What are factors, those can impact the customer will default or not?

- **Income:** Higher Income means higher chance of financial stability that may lead to lower default rate
- **Job Type:** If job is more stable that can also lead to lower default rate. Example: Front End Sales job is less stable than data science job
- **Credit history:** Your previous good repayment behaviour suggest that you know how to use credit products and that leads to lower default rate

Hypothesis generation

Example: What are factors, those can impact the customer will default or not?

- **Income:** Higher Income means higher chance of financial stability that may lead to lower default rate
- **Job Type:** If job is more stable that can also lead to lower default rate. Example: Front End Sales job is less stable than data science job
- **Credit history:** Your previous good repayment behaviour suggest that you know how to use credit products and that leads to lower default rate
- **Education:** Higher education is directly proportionate to awareness about the use of credit products and that may impact the default rate also

Hypothesis generation

Common Question : Should Hypothesis Generation be done before or after looking at the data?

Hypothesis generation to be done before looking at the Data.

Why?

1. Let's you think of all the factors which might affect the problem without being biased.
2. Stops time wastage in analyzing all available data.

Analytics Vidhya

Learn everything about analytics



Thank You

Analytics Vidhya

Learn everything about analytics