

Analyze Credit One Customer Demographic Data to help Credit One to reduce the customer to default on credit obligations

By Shifeng Li

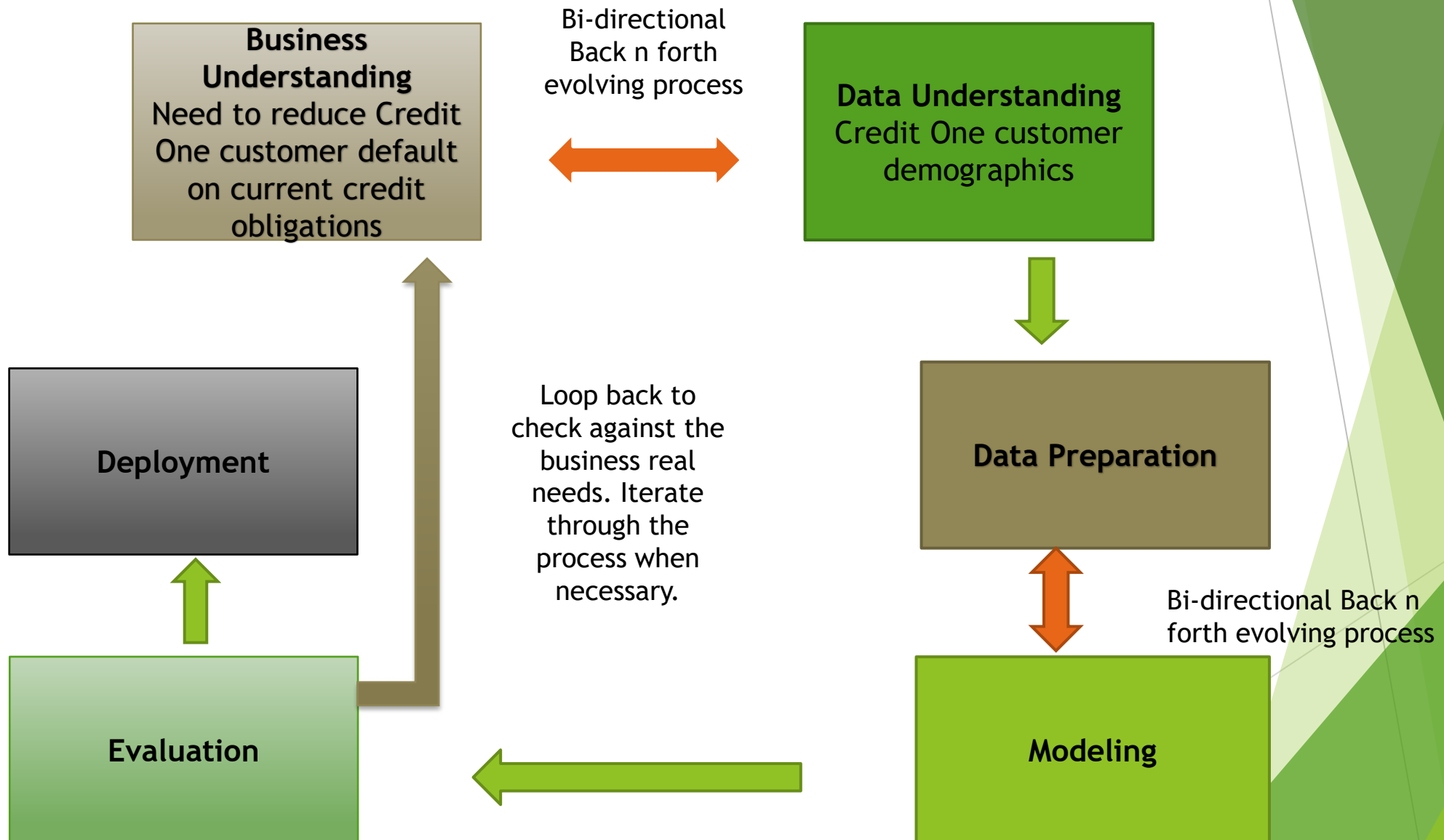
https://github.com/UTOct21DaPtSteve/Steve_Course2.git

In this project, we will conduct comprehensive data science process to analyze through the Credit One customer demographic data, with a goal to identify the core factors that can make the customer likely to default on their current credit obligations.

The goal is find those contributing factors for Credit One to better predict the chance a customer could default on credit debt, and thus properly determine the tailored customized credit level for customers to reduce the risk chance.

The project will systematically go through the data science process steps to gain the good accuracy to identify co-related patterns and to develop prediction model for Credit One.

Iterative Data Science Process Workflow and Cycle



Business Understanding

As the starting point of the Data Science Process, need to gain solid understanding of the Credit One business needs for this project. And develop the business objectives, solutions, and data science process goals for Credit One. Then layout the project plan.

In the past year, Credit One has seen an increase in the numbers of customers have defaulted on loans

Examine the current customer demographics to find out what could contribute to a customer to default on credit obligations

Identify the customized credit level for each customer to reduce the chance the customer would default on loans

Develop prediction model(s) for Credit One to support decision process for future customers.

Data Understanding

The quality of the data used by Data Science Process has the critical impact to the data analysis results. How accurate the discoveries and predication are relying on the quality of the data being analyzed. Therefore, by constantly checking against the business needs and goal, we will go through iterative steps to collect the data, explore the data, and verify the data, to build a solid data foundation for the rest of Data Science Process steps.

Data Preparation

In order to prepare for the data modeling in the Data Science Process, the collected raw data attributes must be properly selected, cleaned, and formatted, to fit into the data modeling tools to process through the computation algorithms.

Therefore, will go through the standard data preparation steps to develop the final data set(s) for modeling process to use.

Modeling

Based on the initial data explore and data analysis, will choose the proper algorithms to experiment through different modeling approaches.

Split the data into training set, test set, and validation set.

Build and run through multiple modeling techniques

Cross compare multiple models data results against the business criteria, and test data. Identify the best model and predication model.

Evaluation

The evaluation step provides the loop back to the Business Understanding Step to provide the iterative opportunity for Data Science Process to meet the business goals at maximum level.

Communicate with Business owner/stakeholder to evaluate how well the data analysis models meet the business goals. In the Data Science Process steps, this step loops back to the starting step, the Business Understanding step, and based on the stakeholder's feedback and acceptance level, will have iterations to go through Data Science Process steps, as needed.

During the evaluation process with the business owners, check if the models meeting business success goals. Check if anything missed? Determine what part can be signed off and deployed, what part still need enhancement. And, to identify new project for more in-depth analysis or bigger scoped multiple domain data analysis.

Deployment

After going through the Data Science Process steps and gain the acceptance from the Credit One business stakeholders, final documentation will be delivered to Credit One as the following components.

The plan to deploy the developed model.

Maintenance plan to support using the model in the Credit One operation.

The Summary report of the delivered model and insights for better solution for the future.