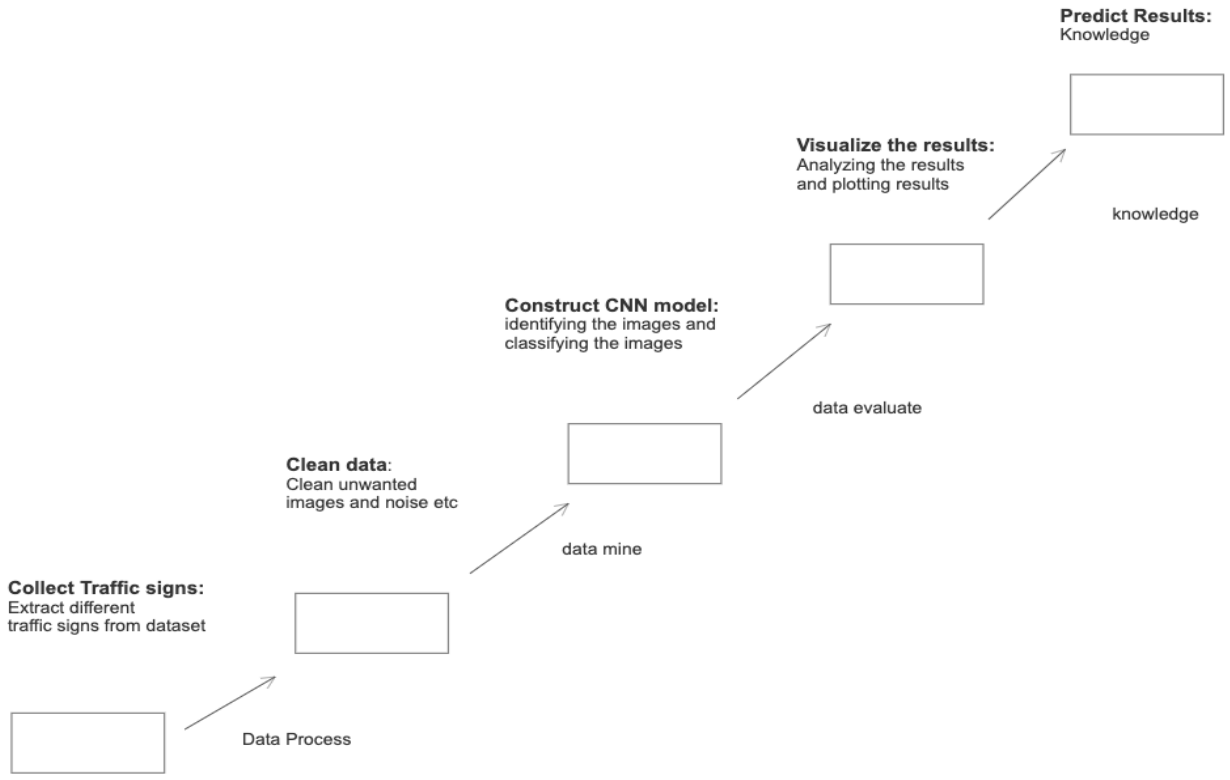
Traffic Sign Recogntion

- **Team**: PirateMiner
- **Team Member**: Sanjay Bhargav Madamanchi

KNOWLEDGE DATA DISCOVERY



Selection:

Selection of raw data from Kaggle

Sources: https://www.kaggle.com/meowmeowmeowmeowmeow/gtsrb-german-traffic-sign

Preprocessing:

Preprocessing the data by removing irrelevant data and dropping the null, duplicate values and clearing off the data which is not in proper format or converting the format is taken care of.

Transformation:

Transforming the data by picking unique values and extracting features and setting up dimensions to make the data efficient.

Data Mining:

Recognizing and identifying the images with the help of CNN model

Interpretation/Evaluation:

Training and testing the model in order to correctly classify the images.

Knowledge Representation:

Visualizing the plots and performing feature analysis and representing the knowledge.

FEATURE ENGINEERING

Features:

Feature Engineering is needed to prepare the input data properly so that our algorithm to determine a favorable candidate can run properly. In the data preprocessing stage of our project, The following are conducted:

1. **Preprocessing the Data:**
   Resizing the images to the required size and converting them into an array and loading the necessary libraries required
2. **Analyzing the data:**
   Dividing the traffic signs present into 43 classes based on the total images present to meet the classifiers
3. **Shuffling the data:**
   The objective of data shuffling to gather and blend up data and derive logical relationships between columns. In other words, this process arbitrarily shuffles the data based on a single or a set of attributes.

4. **Finding the model suitable for the data:**

   Adding various layers in order to gather the features from the image using trial and error method to find the accurate model

5. **Augmenting the data**

   Creating new training data from the existing data . This technique is used to increase the size of the training data to be given as input in the training model by performing minor modifications in the existing data. The Neural network will consider all the images as distinct images which will:

   a).Increase the overall performance of the model

   b).protect overfitting.

6. **Classification:**

   Classifying the data using various metrics to match the required specifications and dividing into the previous 43 classes