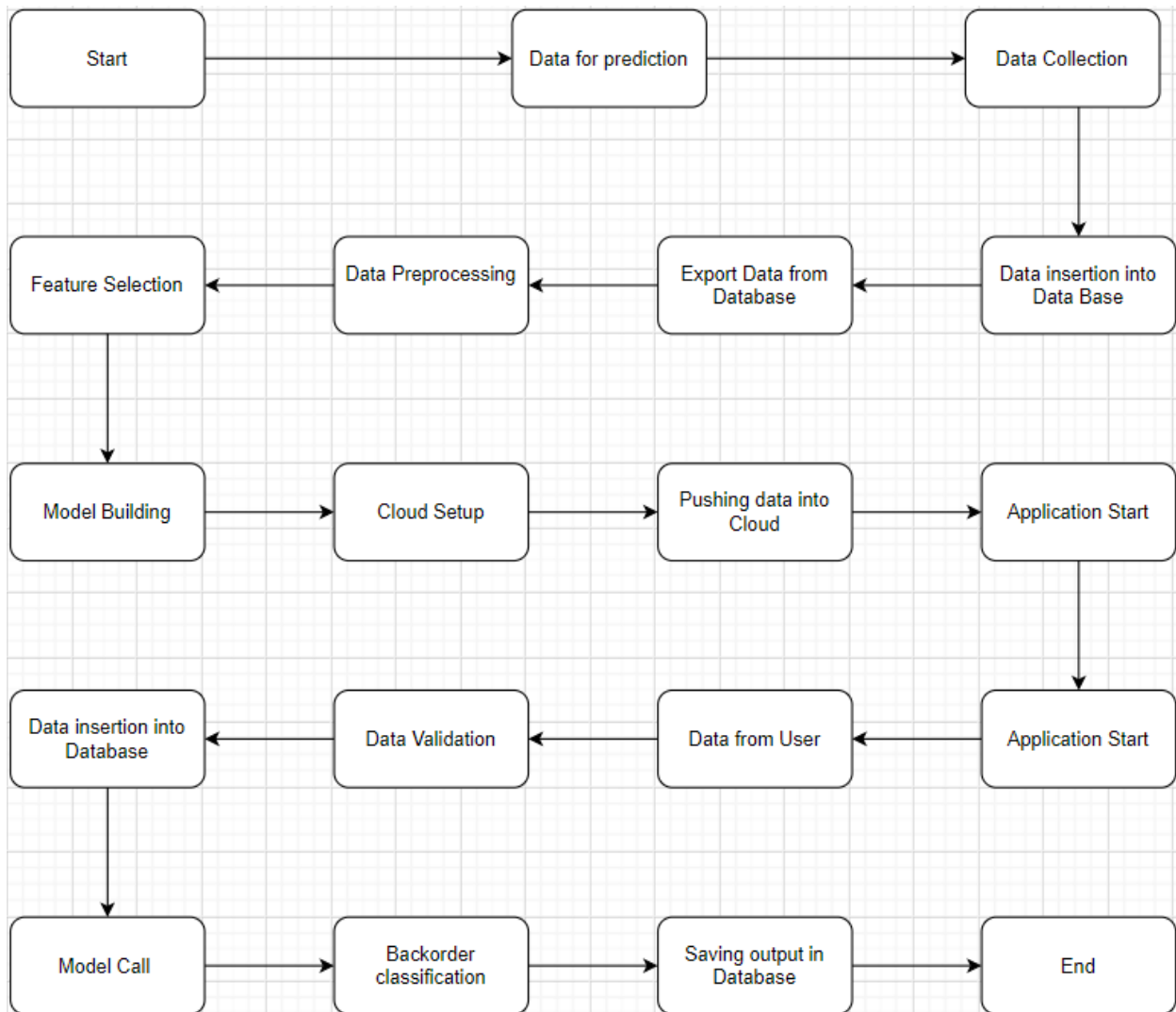


ARCHITECTURE DESIGN  
BACKORDER PREDICTION

## Architecture



## 1. Architecture Description

### 1.1 Data Description

Data consist of inventory details, sales details of past 6 months, forecasted sales of 6 months and backorder details of product. This dataset contains 1048575 rows and 23 columns, where went on backorder is the target variable that contains YES/NO.

### 1.2 Data Collection

Data is available in User system but the data requires some kind of pre-processing techniques to be applied on it.

### 1.3 Data Insertion into Data base

- a) Data Base creation and connection – Create a database with name passed. If the database is already created, open the connection to the database.
- b) Table creation in the database
- c) Insertion of files in the tables

### 1.4 Export Data from Database

Data Export from Database – The data in a stored database is exported as a CSV file to be used for Data Pre-processing and Model Training.

### 1.5 Data Pre-processing

Data Pre-processing steps we could use are Null Value Handling, Label encoding, handling columns that are not Normally Distributed.

### 1.6 Feature Selection

Applying Feature selection techniques for categorical columns and numerical columns to remove features that are not correlated with target variable and feature that are correlated with independent variable itself.

## 1.7 Model Building

After Data preprocessing and Feature selection we will find the best model. For different different feature selection each algorithm will be passed with the best parameters derived from Grid-Search. We will calculate the AUC scores and precision recall scores and select the model with best score. Model will be saved for use of predicting the backorder.

## 1.8 Data From User

We will collect ERP data from user such as inventory details, product lead time, backorder details, forecasted details.

## 1.9 Data Validation

Here Data Validation will be done, given by the user.

## 1.10 User data inserting into Database

Collecting the data from the user and storing it into the database. The database can be either MongoDB or Cassandra.

## 1.11 Model Call for user

Backorder prediction will be done , this output will be saved in Database and it will be used to show the same output if other users provide the same data.

## 1.12 Deployment

We will deploy the model to AWS.