Low Level Design (LLD)

**US VISA APPROVAL**

**Anna Elsa Luiz**

|  |  |  |  |
| --- | --- | --- | --- |
| Date Issued | Version | Description | Author |
| 22-03-2024 | 1 | Initial LLD -1.0 | ANNA ELSA LUIZ |

Document Version Control

**Contents**

Document Version Control

Abstract

1. Introduction
   1. What is Low-Level Document
   2. Scope
2. Architecture
3. Architecture Description
   1. Data Description
   2. Data Ingestion
   3. Data Export From DataBase
   4. Pre-Processing
   5. Feature Extraction
   6. Model Training
   7. Data From User
   8. Pre-Processing user Data
   9. Prediction
   10. Result on User-Interface
   11. Deployment

**Abstract**

This Low-Level Design (LLD) document focuses on implementing a predictive model for US visa approval, detailing the architectural components, data structures, data flow, and algorithms crucial for deployment. It offers a comprehensive portrayal of the data processing pipeline, covering stages such as data collection, preprocessing, feature engineering, model training, and evaluation.

**Introduction**

**1.1. What is Low-Level design document?**

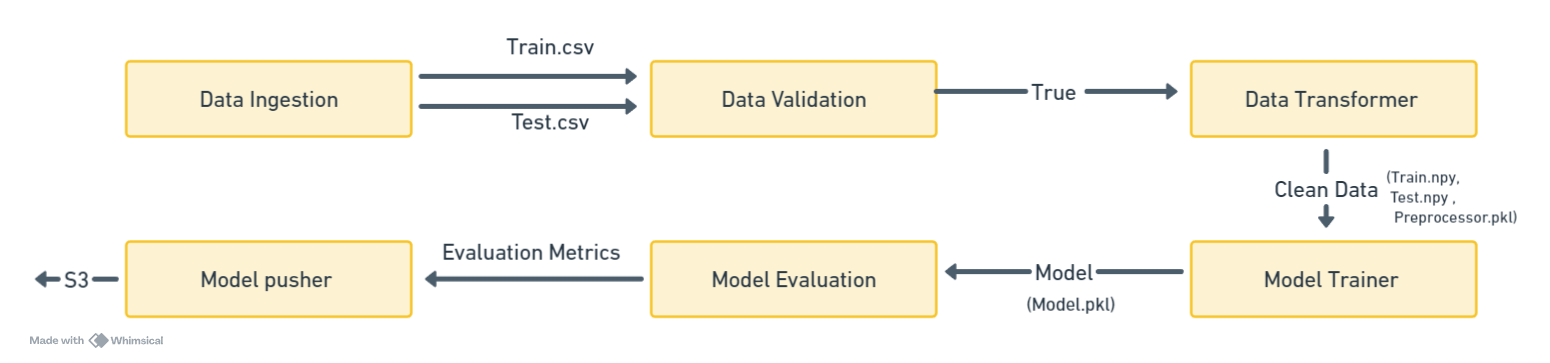
The goal of LLD or a low-level design document (LLDD) is to give the internal logical design of the actual program code for Food Recommendation System. LLD describes the class diagrams with the methods and relations between classes and program specs. It describes the modules so that the programmer can directly code the program from the document.

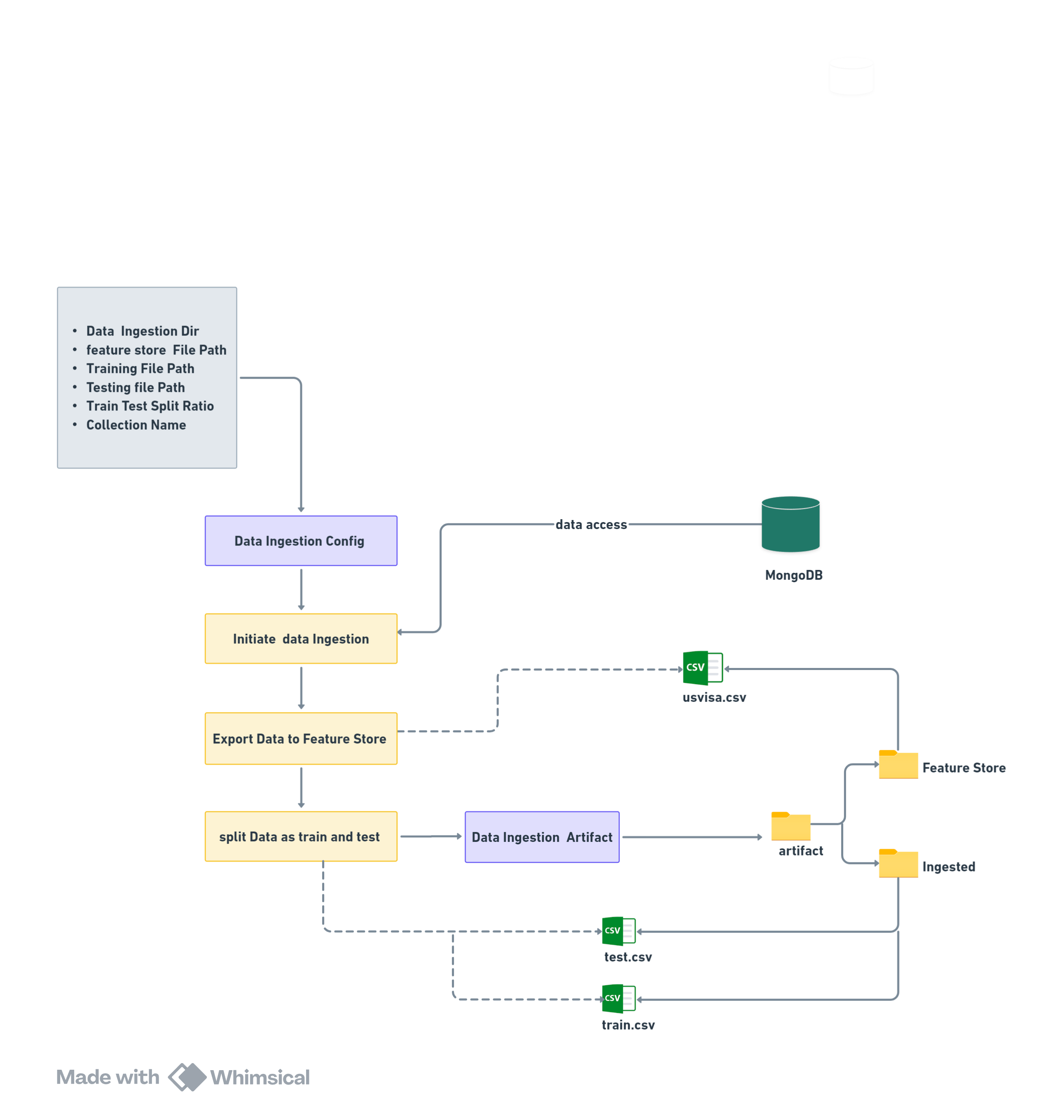
**1.2. Scope**

Low-level design (LLD) is a component-level design process that follows a step-by-step refinement process. This process can be used for designing data structures, required software architecture, source code and ultimately, performance algorithms. Overall, the data organization may be defined during requirement analysis and then refined during data design work.

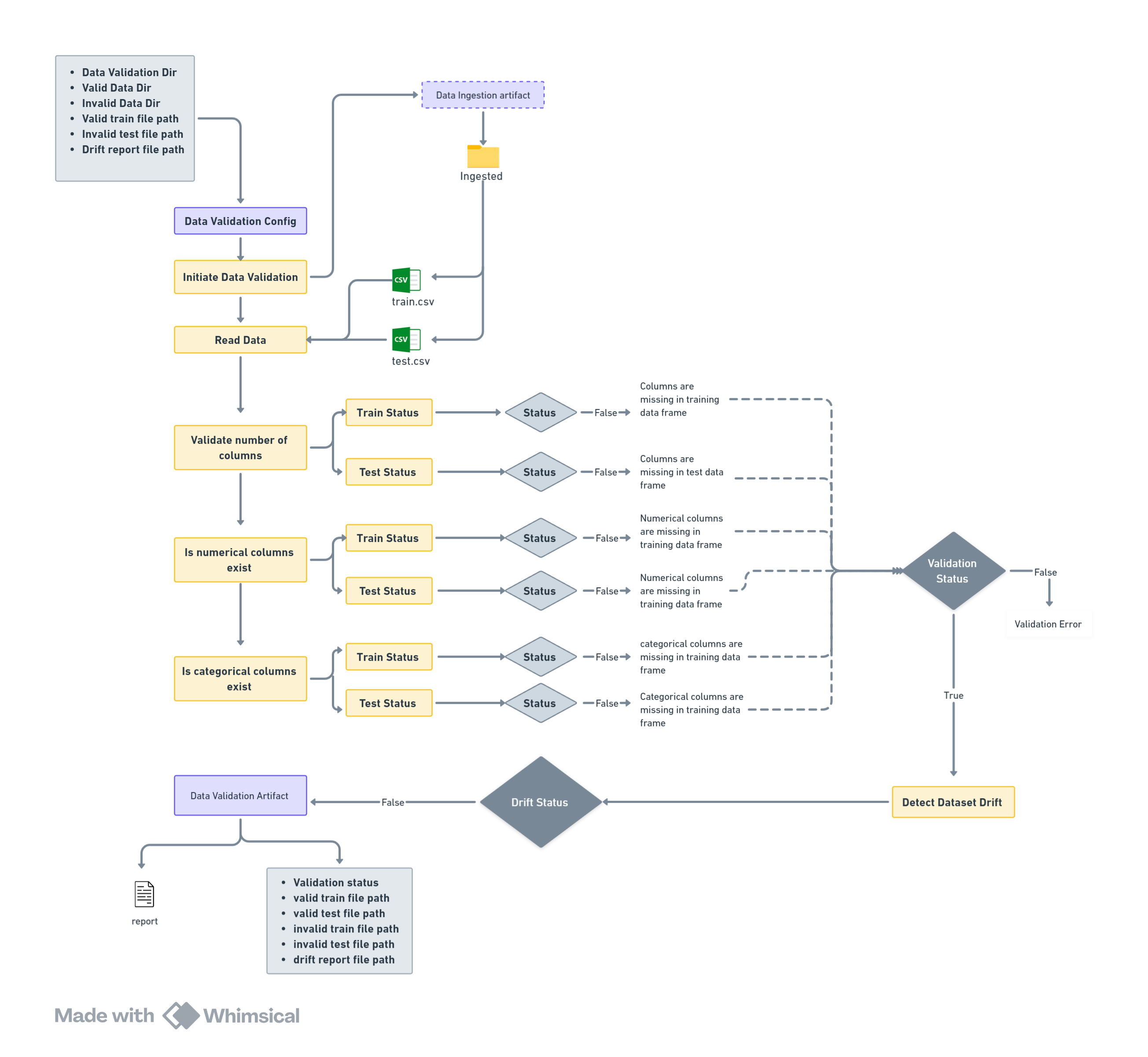
**2. Architecture**

**Pipeline**

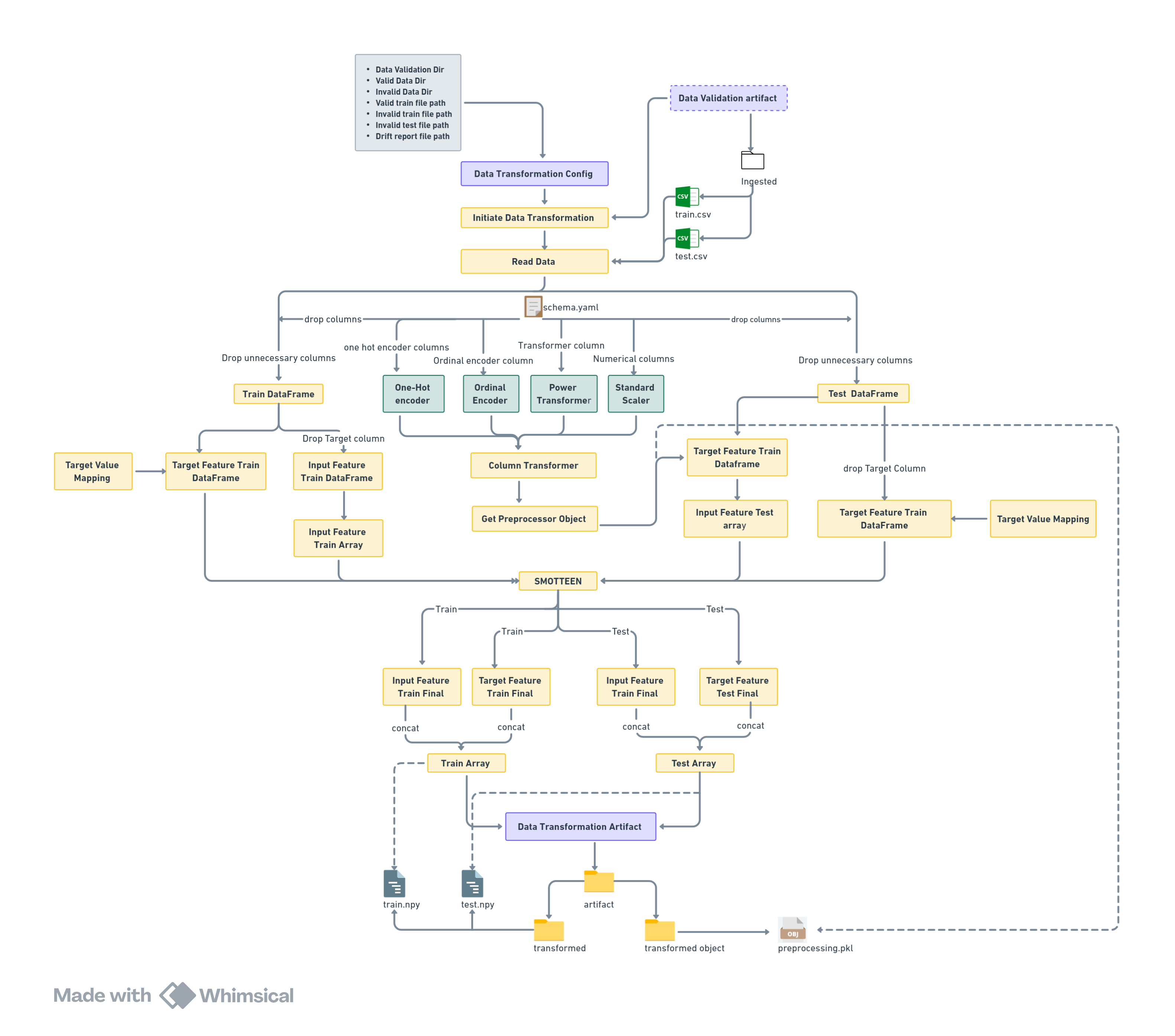


**Data Ingestion** 

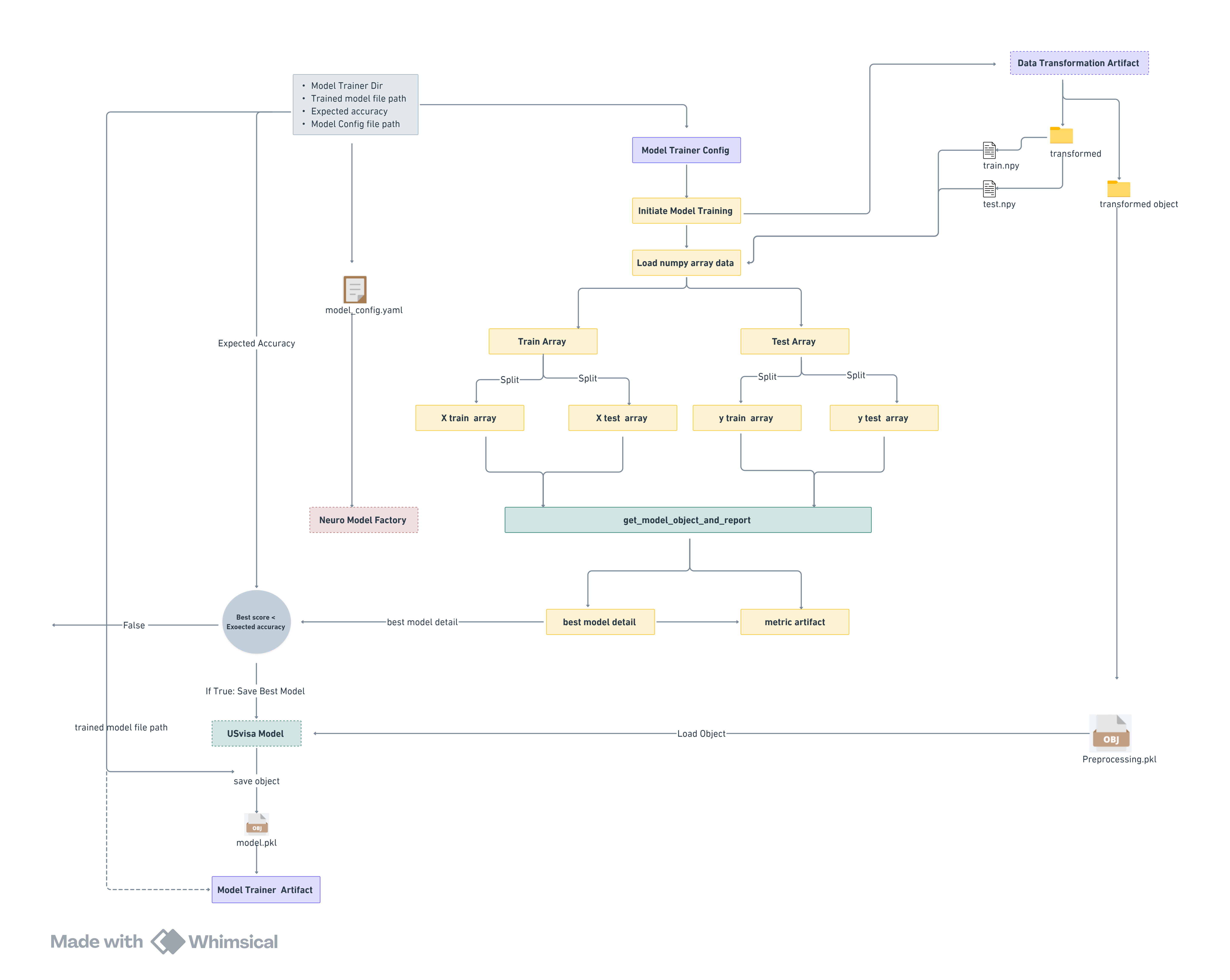
**Data validation**

****

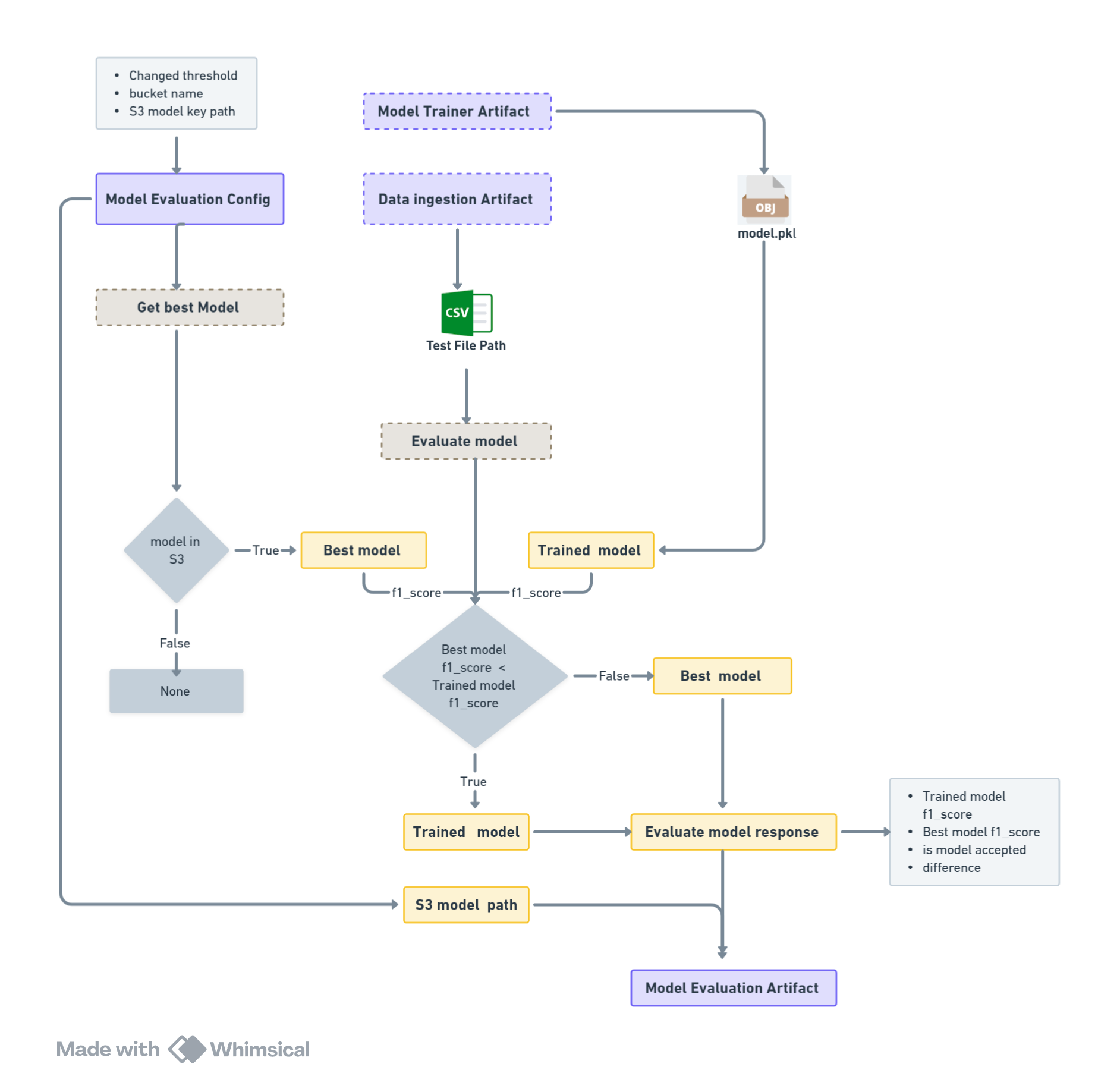
**Data Transformation**

****

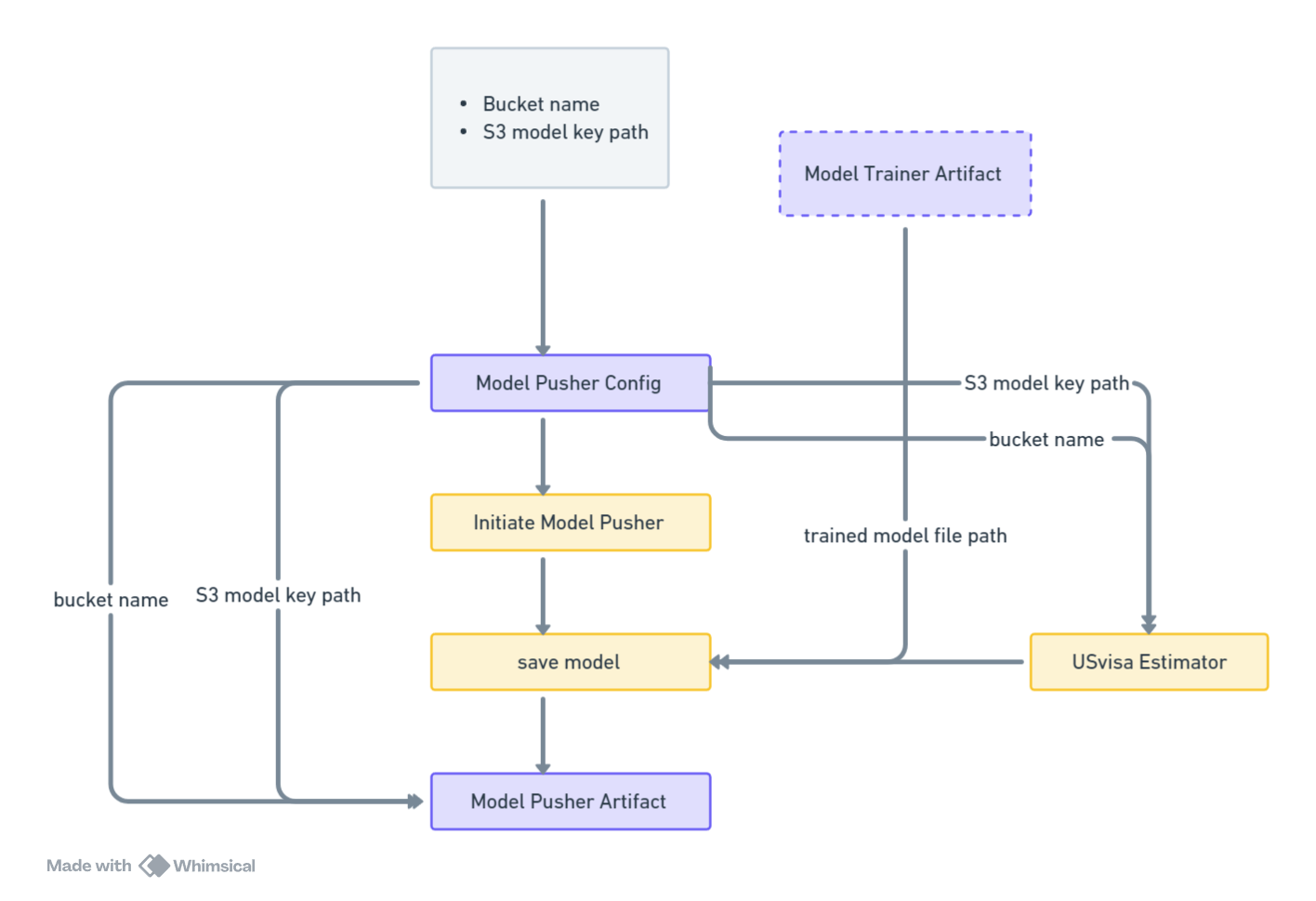
**Model Trainer**

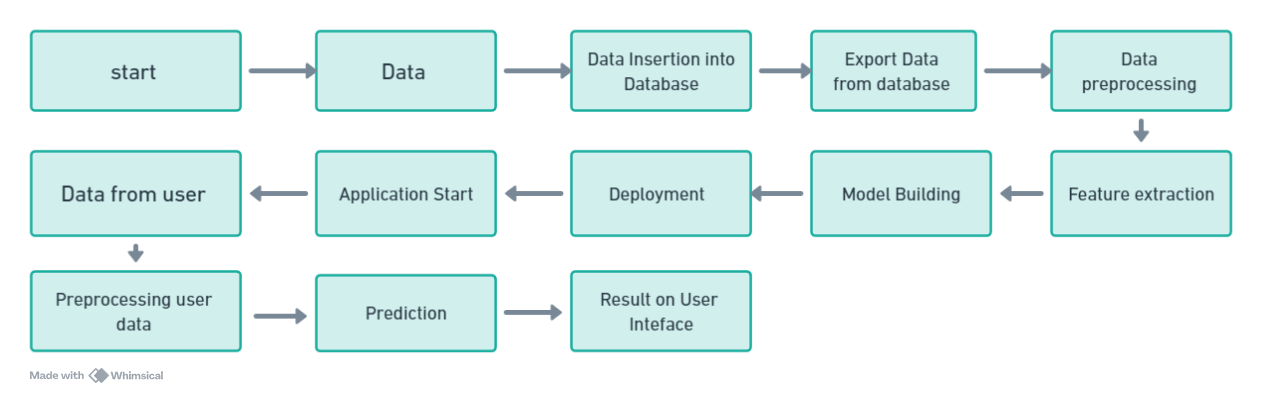
****

**Model Evaluation**

****

**Model Pusher**

****



**2.1 Data Description**

The Dataset is part of Office of Foreign Labor Certification (OFLC). The data consists of 25480 Rows and 12 Columns

* case\_id: ID of each visa application
* continent: Information of continent the employee
* education\_of\_employee: Information of education of the employee
* has\_job\_experience: Does the employee has any job experience? Y= Yes; N = No
* requires\_job\_training: Does the employee require any job training? Y = Yes; N = No
* no\_of\_employees: Number of employees in the employer's company
* yr\_of\_estab: Year in which the employer's company was established
* region\_of\_employment: Information of foreign worker's intended region of employment in the US.
* prevailing\_wage: Average wage paid to similarly employed workers in a specific occupation in the area of intended employment. The purpose of the prevailing wage is to ensure that the foreign worker is not underpaid compared to other workers offering the same or similar service in the same area of employment.
* unit\_of\_wage: Unit of prevailing wage. Values include Hourly, Weekly, Monthly, and Yearly.
* full\_time\_position: Is the position of work full-time? Y = Full Time Position; N = Part Time Position
* case\_status: Flag indicating if the Visa was certified or denied

**2.2 Data Ingestion**

The database employed for data storage in this project is MongoDB. As part of the database setup process, a database is created with the specified name. If the database already exists, the system establishes a connection to it. Subsequently, tables are created within the database to organize the stored data effectively. Finally, files are inserted into these tables, ensuring that the data is appropriately structured and accessible for subsequent operations and analysis.

**2.3 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.

**2.4 Data Pre-Processing**

Data preprocessing refers to the steps taken to transform raw data into a format that is suitable for analysis and modelling.

**2.5 Feature Extraction:**

Feature extraction is the process of selecting and transforming relevant data characteristics to represent and capture meaningful information.

**2.6 Model Building:**

Model Building refers to the process of developing a machine learning model that can learn patterns from data and make predictions or perform tasks based on those patterns.

**2.7 Data From User**

Here we will collect the data from the user. User will directly Input the Data from User-Interface. After that data will converted as dataframe and stored.

**2.8 Pre-Processing User Data**

This involves doing pre-processing steps on user-input data like Imputing Null values, Transforming categorical data to numerical data and Scaling the Numerical data.

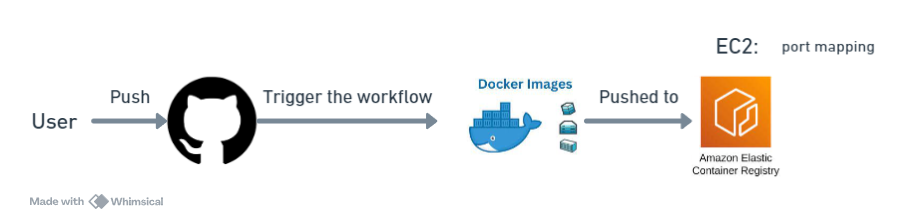
**2.9 Prediction**

In this step, using the best-performing model and predicted the output based on pre-processed user-input data and.

**2.10 Result on User-Interface**

In the last the Result will be shown on the UI with the help of HTML

**2.11 Deployment**

****