

CP353 : Systems Analysis and Design

โปรแกรมการวิเคราะห์และทำเหมืองข้อมูล

MEMBER



นายกฤตณ์ ชจรกุลญาณ
65102010190



นายพัฒนพงศ์ สุริยวงศ์ปีร์
65102010123



นายบูรพา ยืนยง
65102010418

SYSTEM REQUEST

1. Tangible Value : สามารถลดเวลาในการทำ data mining ลง 20%
2. Intangible Value : สามารถจัดการข้อมูลได้อย่างง่าย และสะดวกมากยิ่งขึ้น

REQUIREMENT DEFINITIONS

FUNCTIONAL REQUIREMENTS

- การนำเข้าข้อมูล : รองรับการนำเข้าข้อมูลจากไฟล์ CSV, Excel,ฐานข้อมูล (SQL, NoSQL), หรือ API
- การจัดการข้อมูล : สามารถเพิ่ม, ลบ, แก้ไขข้อมูลได้
- การคำนวณและแสดงผล : คำนวณและแสดงผลต่างๆ เช่น ค่าเฉลี่ย, ส่วนเบี่ยงเบนมาตรฐาน
- การแปลงข้อมูล : เปลี่ยนแปลงรูปแบบข้อมูล, การแปลงข้อมูลเชิงหมายความที่เป็นเชิงตัวเลข
- การสร้างภาพข้อมูล : สร้างกราฟ, แผนภูมิ เช่น Histogram, Scatter/Box Plot
- การวิเคราะห์สถิติ : คำนวณค่าสถิติต่างๆ เช่น ค่าเฉลี่ย, ส่วนเบี่ยงเบนมาตรฐาน

REQUIREMENT DEFINITIONS

FUNCTIONAL REQUIREMENTS

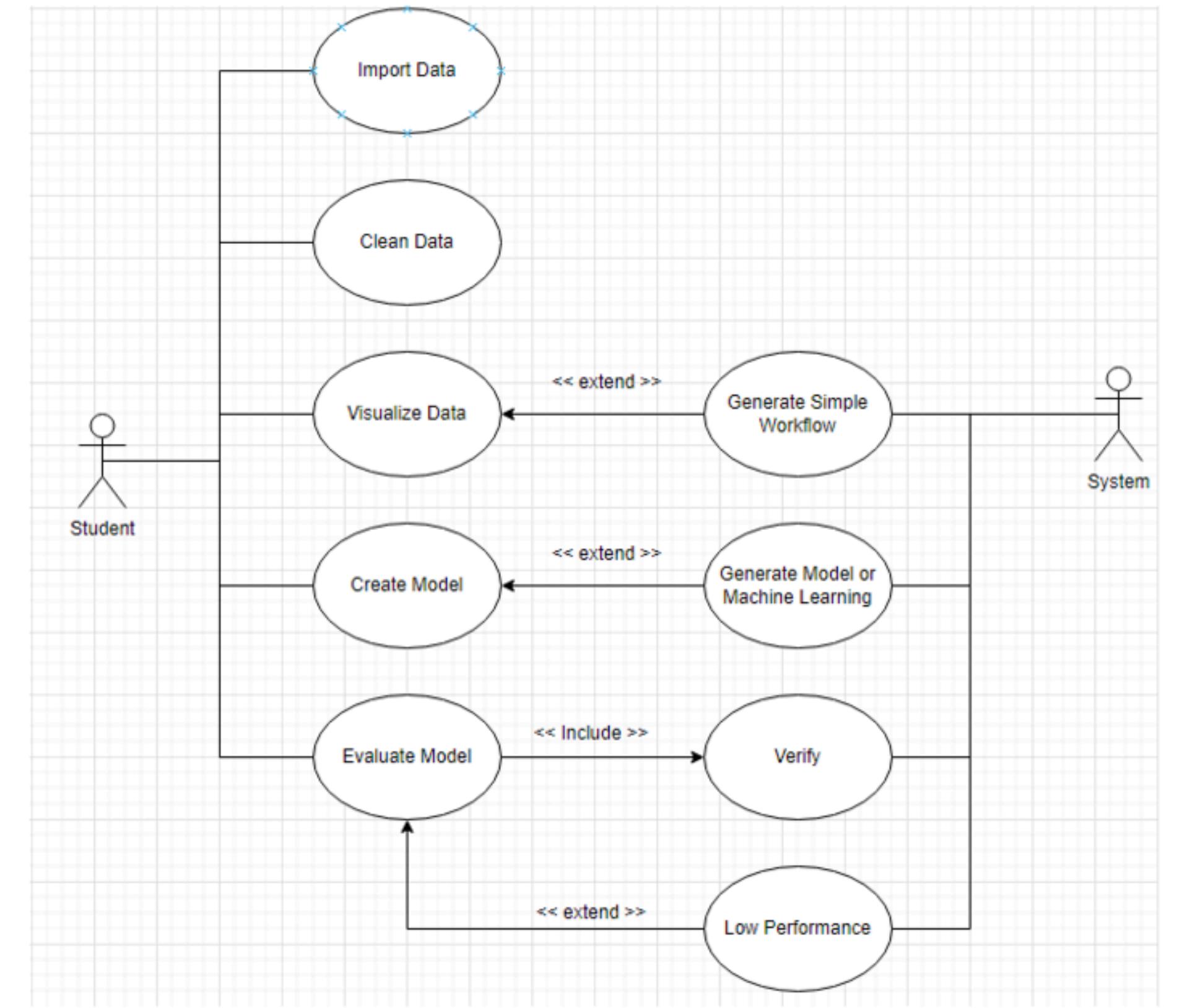
- การเลือก Algorithm : มี Algorithm ให้เลือกหลากหลาย เช่น Linear Regression, Decision Tree, Random Forest, Neural Network
- การฝึกสอนแบบจำลอง: สามารถฝึกสอนแบบจำลองได้จ่าย
- การประเมินผลแบบจำลอง : คำนวณค่าประเมินผล เช่น Accuracy, Precision, Recall, F1-Score
- การบันทึกและโหลดแบบจำลอง : สามารถบันทึกและโหลดแบบจำลองที่สร้างไว้
- การใช้งานแบบจำลอง : สามารถนำแบบจำลองที่สร้างไว้ไปทำงานค่าใหม่

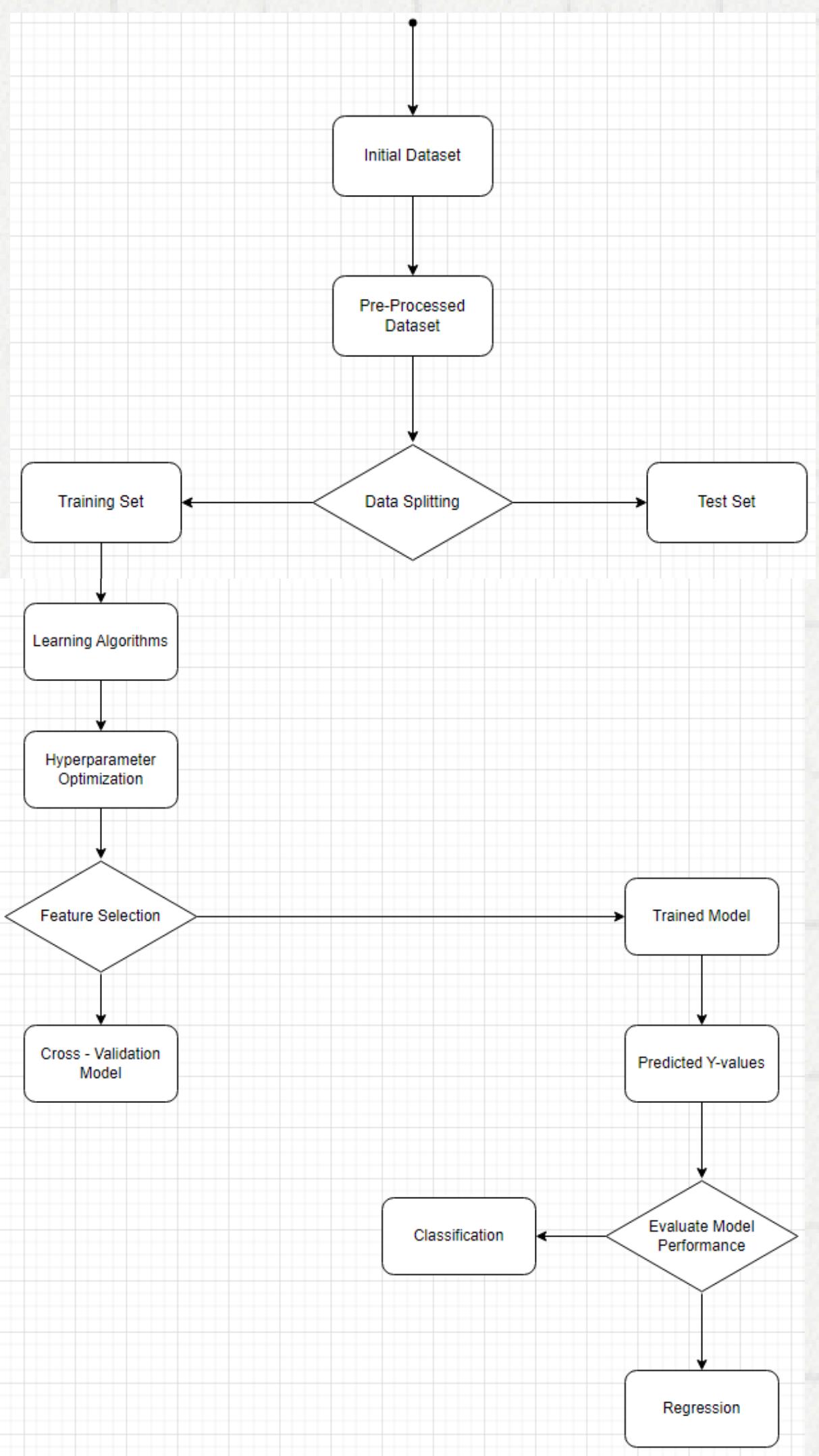
REQUIREMENT DEFINITIONS

NON-FUNCTIONAL REQUIREMENTS

- ประสิทธิภาพ: ระบบต้องทำงานได้รวดเร็ว และมีกับข้อมูลขนาดใหญ่
- ความสามารถในการปรับขยาย: ระบบสามารถรองรับการเพิ่มปริมาณข้อมูลและผู้ใช้งานได้
- ความปลอดภัย: ข้อมูลของผู้ใช้งานต้องได้รับการรักษาความปลอดภัย
- ความเสถียร: ระบบต้องทำงานได้อย่างต่อเนื่อง ไม่ค่อยมีปัญหา
- ความสามารถในการใช้งาน: ใช้งานง่าย มีอินเทอร์เฟซที่เป็นมิตรความเข้ากันได้: รองรับการทำงานบนระบบปฏิบัติการและเบราว์เซอร์ต่างๆ

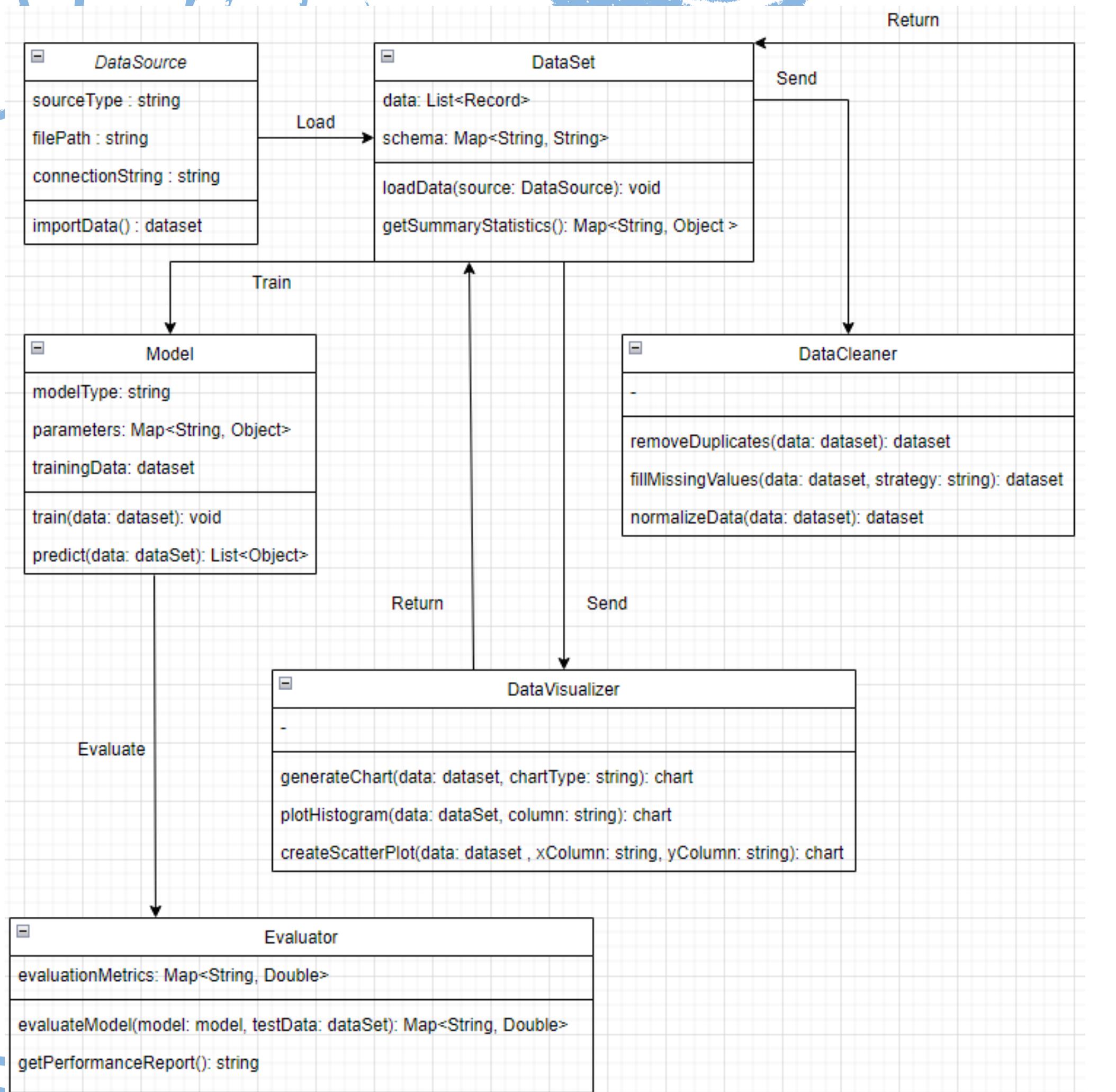
Use-Case Diagram



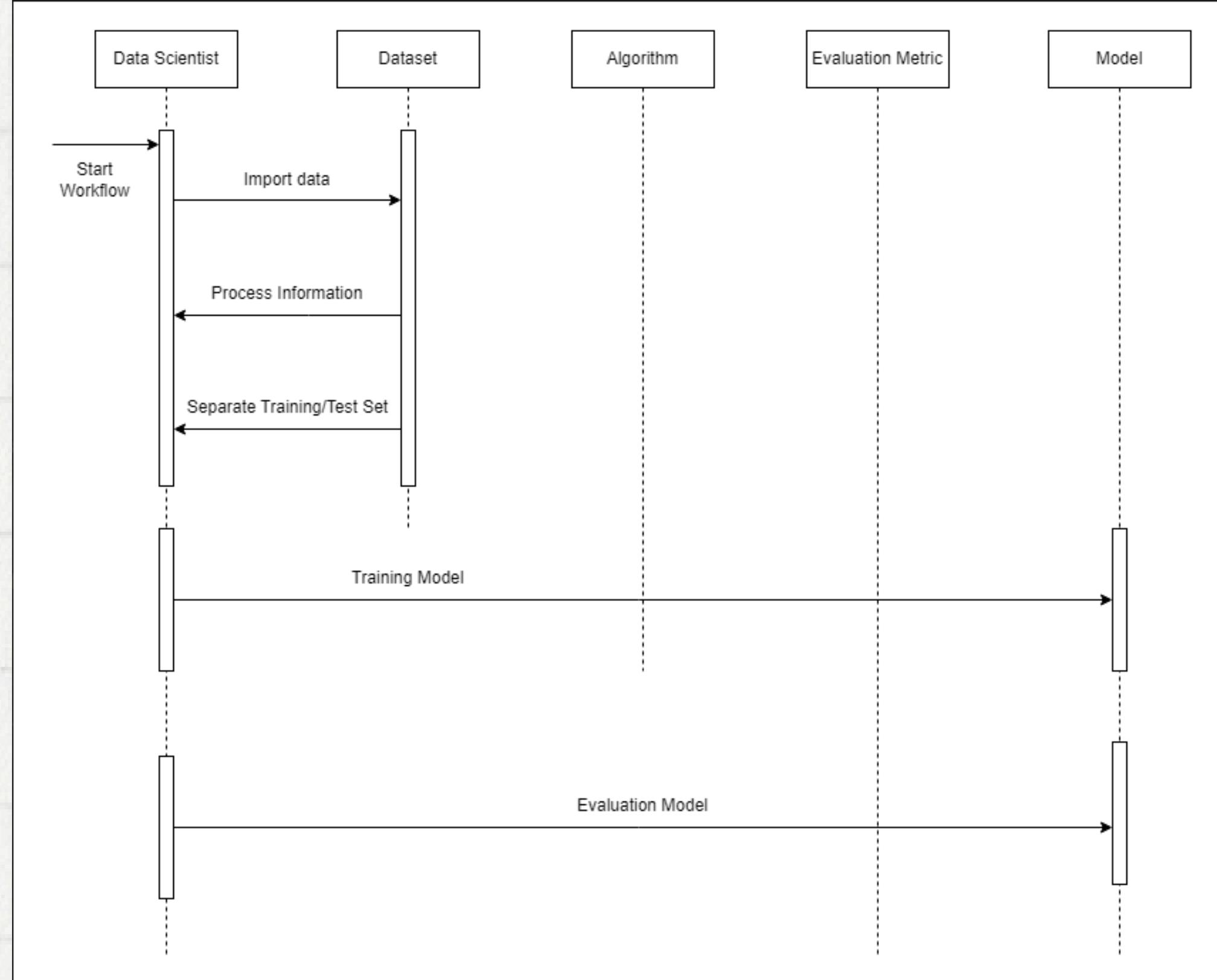


Activity Diagram

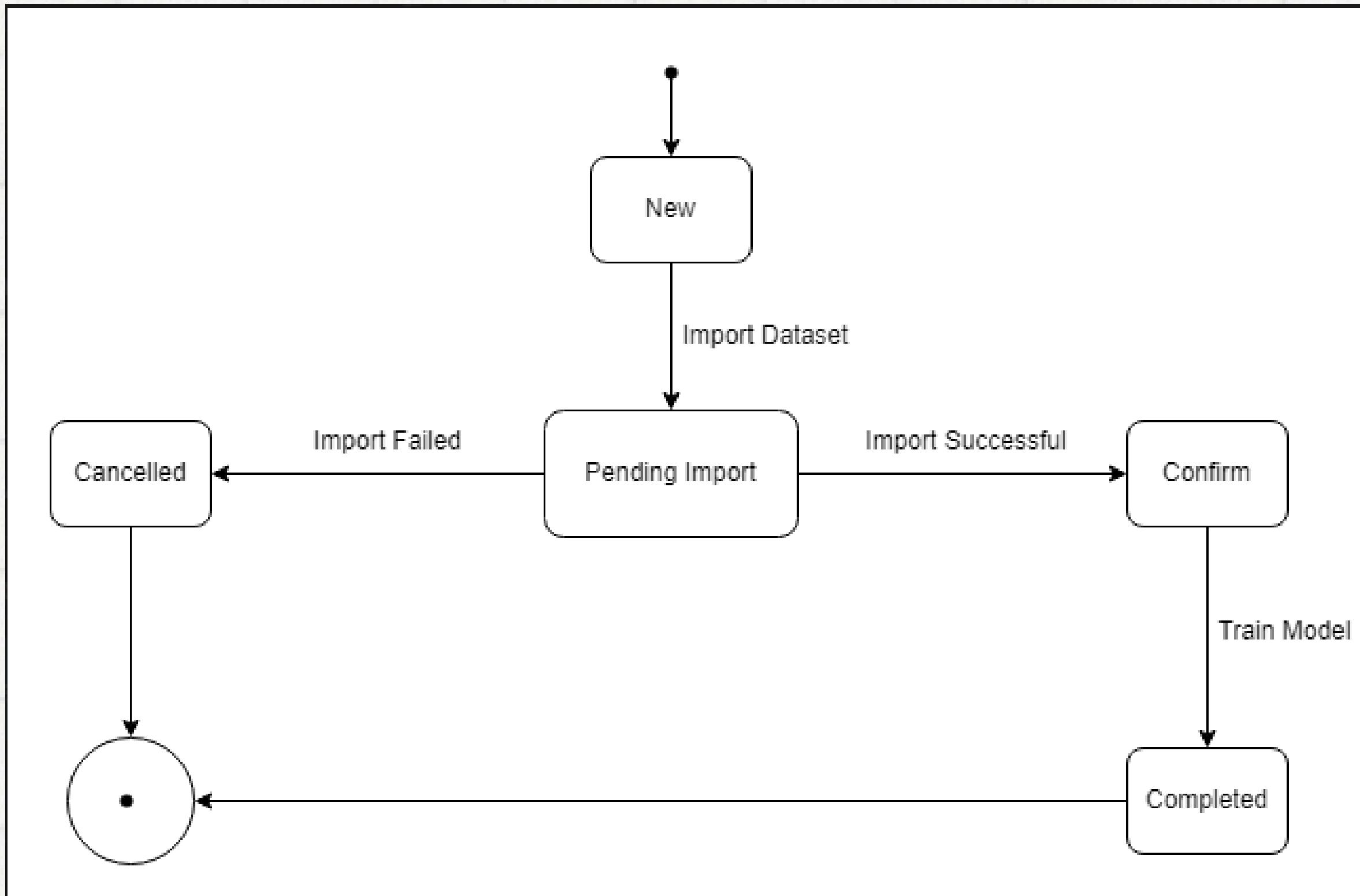
Analysis Class Diagram



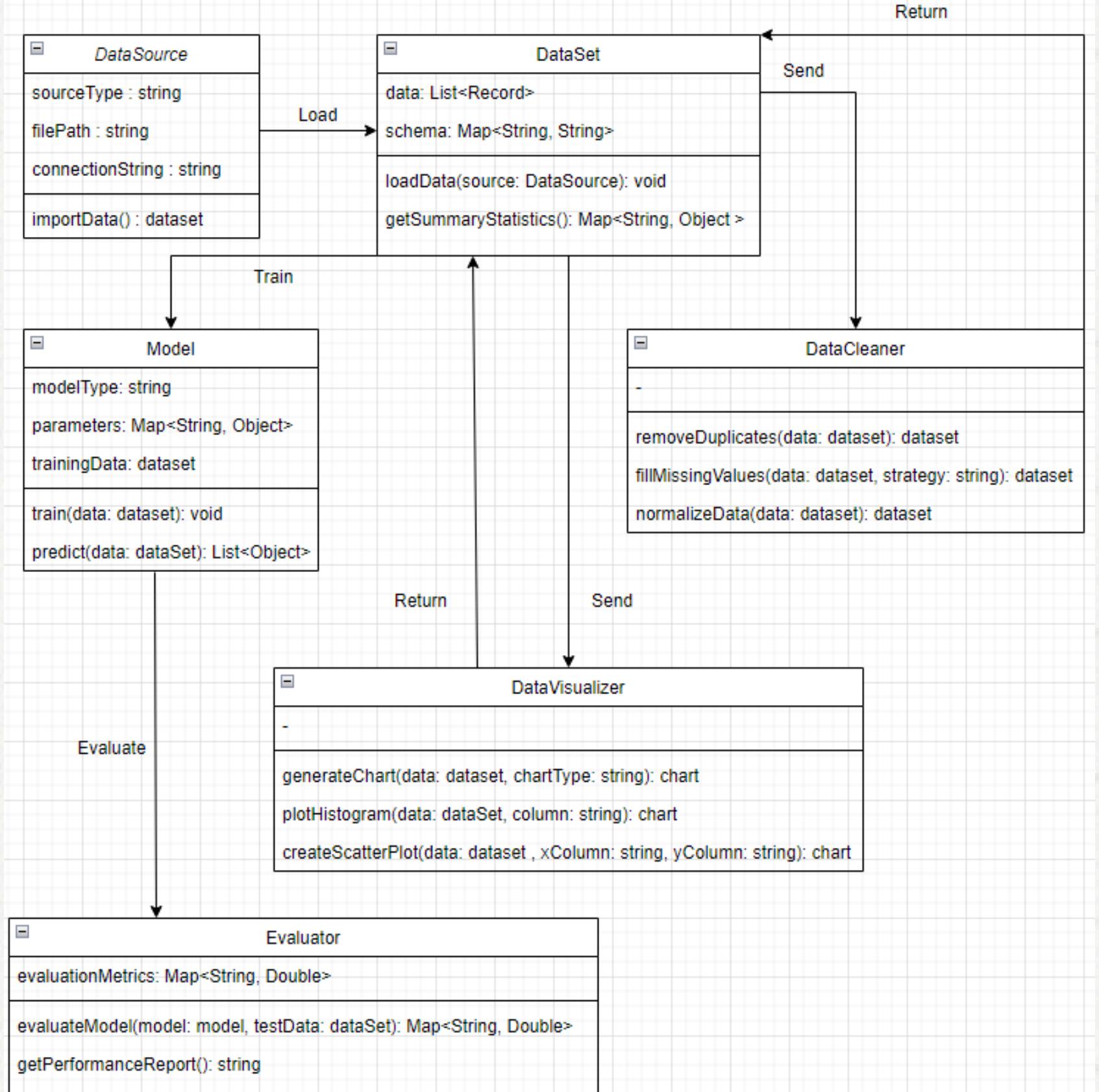
Sequence & Communication Diagram



Behavioral State Machine



Design Class Diagram



การเพิ่ม Attributes กี่จำเป็นลงในคลาส DataSet :

- การเพิ่มการจัดเก็บข้อมูลชั่วคราว (Cache)
- การเพิ่ม Metadata ของข้อมูล

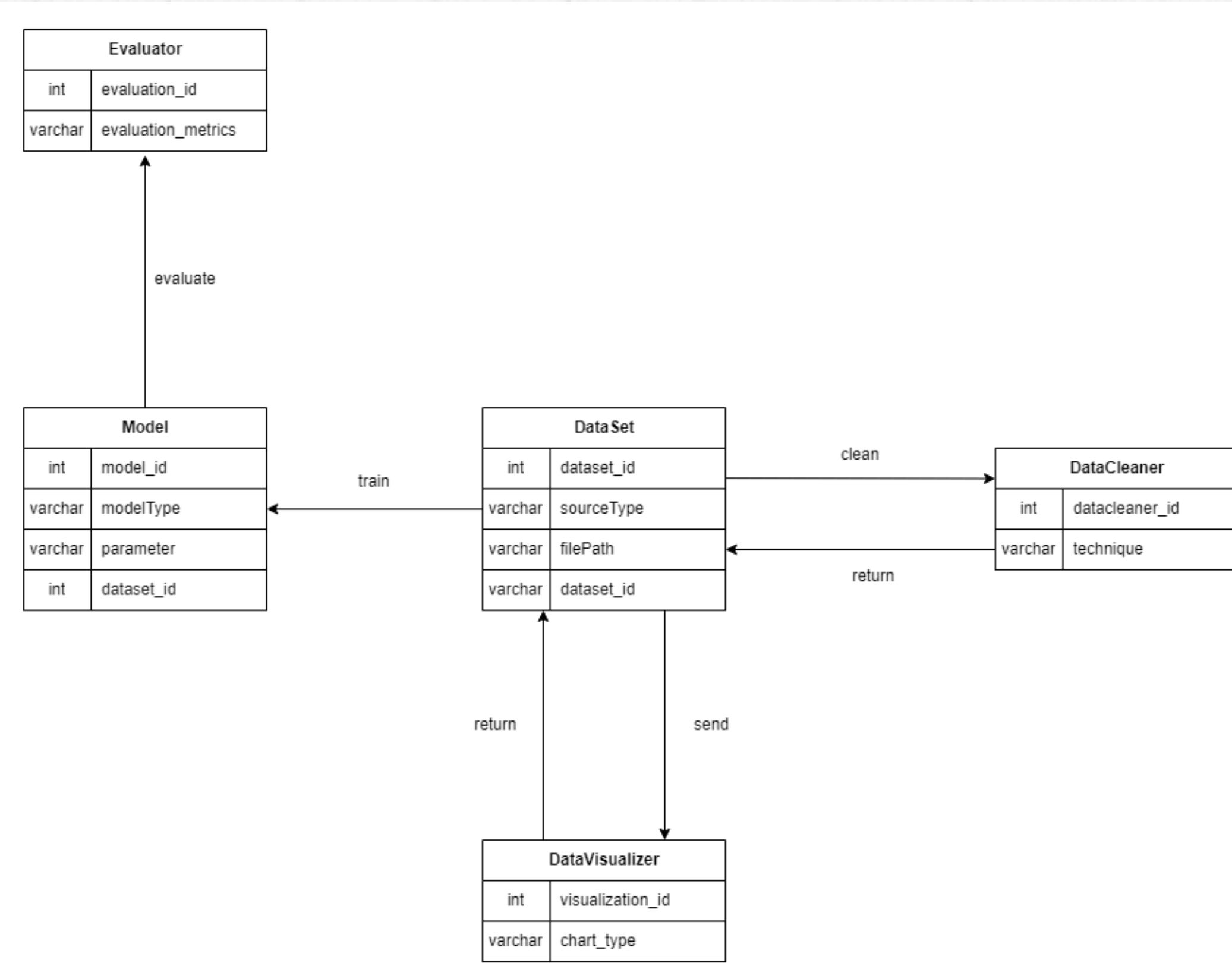
Method Contract

method name: getSummaryStatistics	class name: DataSet	ID: 1
Clients(costumer): ผู้ใช้งานที่ต้องการตรวจสอบค่าทางสถิติของข้อมูลใน DataSet		
Associated Use Cases: Visualize Data , Analyze Data		
Description of Responsibilities: ดึงข้อมูลสถิติพื้นฐาน (mean, max, min, standard deviation) ของแต่ละคอลัมน์ จาก DataSet และคืนค่าในรูปแบบ Map หรือ JSON		
Arguments Received: datasetID: int		
Type of Value Returned: Map<String, Map<String, Object>> หรือ JSON (key เป็นชื่อคอลัมน์ , value เป็นสถิติ เช่น mean , max , min , std_dev)		
Pre-conditions: <ul style="list-style-type: none">- datasetID ต้องมีอยู่ในระบบ (datasetID != null และ datasetID มีในตาราง DataSet)- DataSet ต้องไม่ว่างเปล่า (ต้องมีข้อมูล)		
Post-conditions: หากการดึงข้อมูล - ประมวลผลสำเร็จ: คืนค่า Map หรือ JSON ที่มีข้อมูลสถิติของแต่ละคอลัมน์ หากการดึงข้อมูล - ประมวลผลล้มเหลว: คืนค่า null หรือ error message หาก datasetID ไม่มีในระบบ หรือไม่มีข้อมูลใน DataSet		

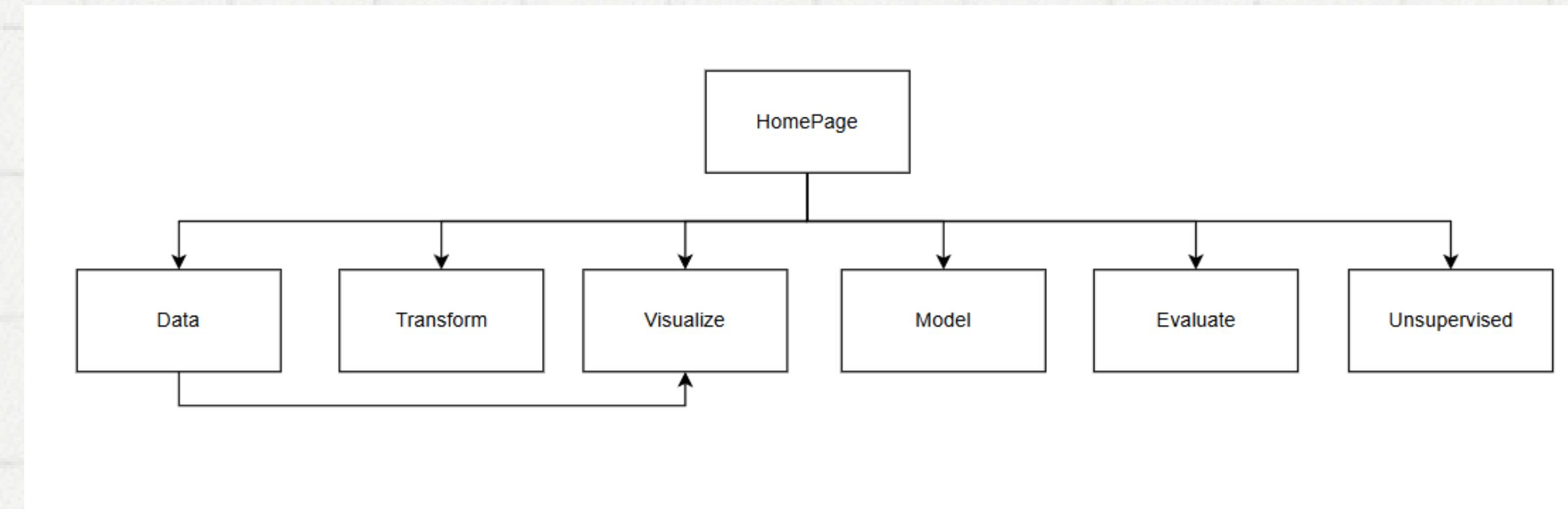
Method Specification

method name: getSummaryStatistics	class name: DataSet	ID: 1		
Trigger/Event: A user requests statistical values (mean, max, min, standard deviation) from a dataset in the system.				
Arguments Received:		Notes:		
datasetID	int	unique ID of a dataset		
Message Sent & Arguments Passed: DataSet.getSummaryStatistics()	Map<String, Object> or JSON	based on the dataset ID		
ClassName.MethodName: DataSet.getSummaryStatistics				
Type of Value Returned: Map<String, Object> or JSON	Contains statistical data of the dataset			
Algorithm Specification: <ol style="list-style-type: none">1) Validate Input:<ul style="list-style-type: none">- Ensure datasetID exists in the system (datasetID != null) and is linked to a valid dataset.- If not valid, return an error message indicating the issue.2) Process Data:<ul style="list-style-type: none">- Retrieve the dataset associated with datasetID.- Compute the following for each column: Mean , Max , Min , Standard Deviation3) Output Result: Return the computed data as a Map or JSON object.				
Misc.Notes: This method assumes that the dataset contains numeric columns for valid statistical calculations.				

Database



Use Interface Design: Windows Navigation Diagram



**Thank you
very much!**