

Low-Level Design

API Endpoints

1. **POST /train-predict**

Description: Trains a model using provided training data and predicts results on test data.

Field	Details
Method	POST
Content-Type	multipart/form-data
Form Fields	trainData (CSV), testData (CSV)
Response	{ "runId": "string", "status": "success" }
Errors	400 (Bad Request), 500 (Internal Error)
Auth Required	No

2. **POST /models/predict**

Description: Runs prediction using an existing trained model.

Field	Details
Method	POST
Content-Type	multipart/form-data
Form Field	file (CSV)
Response	JSON predictions
Errors	400, 500
Auth Required	No

3. **POST /clean**

Description: Cleans the uploaded dataset.

Field	Details
Method	POST
Content-Type	multipart/form-data
Form Field	file (CSV)
Response	Cleaned file (as blob/download)
Errors	400, 500
Auth Required	No

4. GET /runs/:runId/artifacts/figures

Description: Lists generated figure files for a run.

Field	Details
Method	GET
Path Params	runId
Response	[{ "path": "filename.png", "url": "/api/runs/<runId>/artifacts/figures/filename.png" }]
Errors	404
Auth Required	No

5. GET /runs/:runId/artifacts/figures/:filename

Description: Serves an individual figure image.

Field	Details
Method	GET
Path Params	runId , filename
Response	image/png or image/jpeg
Errors	400 (Invalid type), 404
Auth Required	No

6. GET /runs/:runId/artifacts/features-dropped

Description: Returns plain-text list of features dropped during model training.

Field	Details
Method	GET
Path Params	<code>runId</code>
Response	Plain text (feature list)
Errors	404
Auth Required	No

7. `GET /runs/:runId/artifacts/predictions/confusion_matrix.png`

Description: Serves confusion matrix image from prediction results.

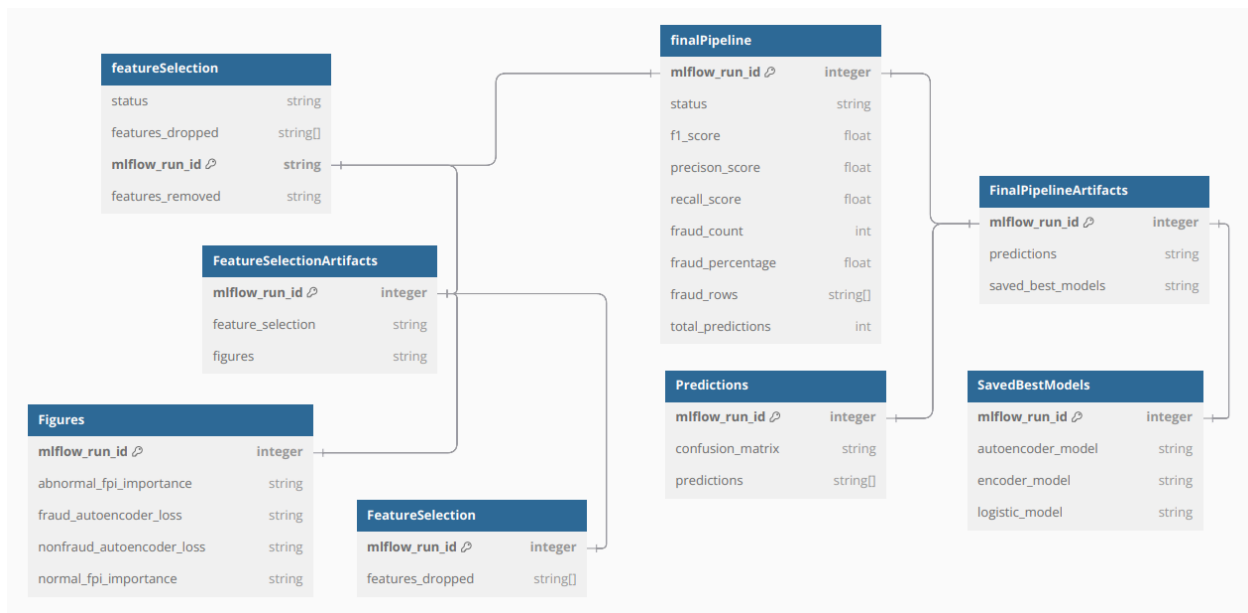
Field	Details
Method	GET
Path Params	<code>runId</code>
Response	<code>image/png</code>
Errors	404
Auth Required	No

8. `GET /runs/:runId/metrics`

Description: Returns model performance metrics in JSON.

Field	Details
Method	GET
Path Params	<code>runId</code>
Response	JSON: Accuracy, Precision, etc.
Errors	404 (Missing file), 500 (Parse error)
Auth Required	No

Data Models



Tools & Technologies

API Framework	Express.js
ML Tracking	MLflow
Data Versioning	DVC
Storage	MongoDB + Local Filesystem
Containerization	Docker

Example Workflow

Scenario: User uploads data and views results

1. `POST /train-predict` → Returns `runId: "abc123"`
2. Poll `GET /runs/abc123/metrics` until `status ≠ "running"`
3. `GET /runs/abc123/artifacts/figures` → List available plots
4. `GET /runs/abc123/artifacts/predictions/confusion_matrix.png` → Render image

Appendix

- **Database Diagram:** [Attach your existing diagram here]

- **Sample Curl Requests:**

```
bash
```

```
curl -X POST -F "trainData=@data.csv" http://localhost:5000/train-predict
```