Machine Learning Forecasting System with AWS Lambda

This solution predicts monthly spend volumes exceeding €100,000 using an engineered XGBoost ML model with excellent accuracy.

The system is fully automated with AWS Lambda, EventBridge, and S3 event notifications.

Performance Metrics

€9.01
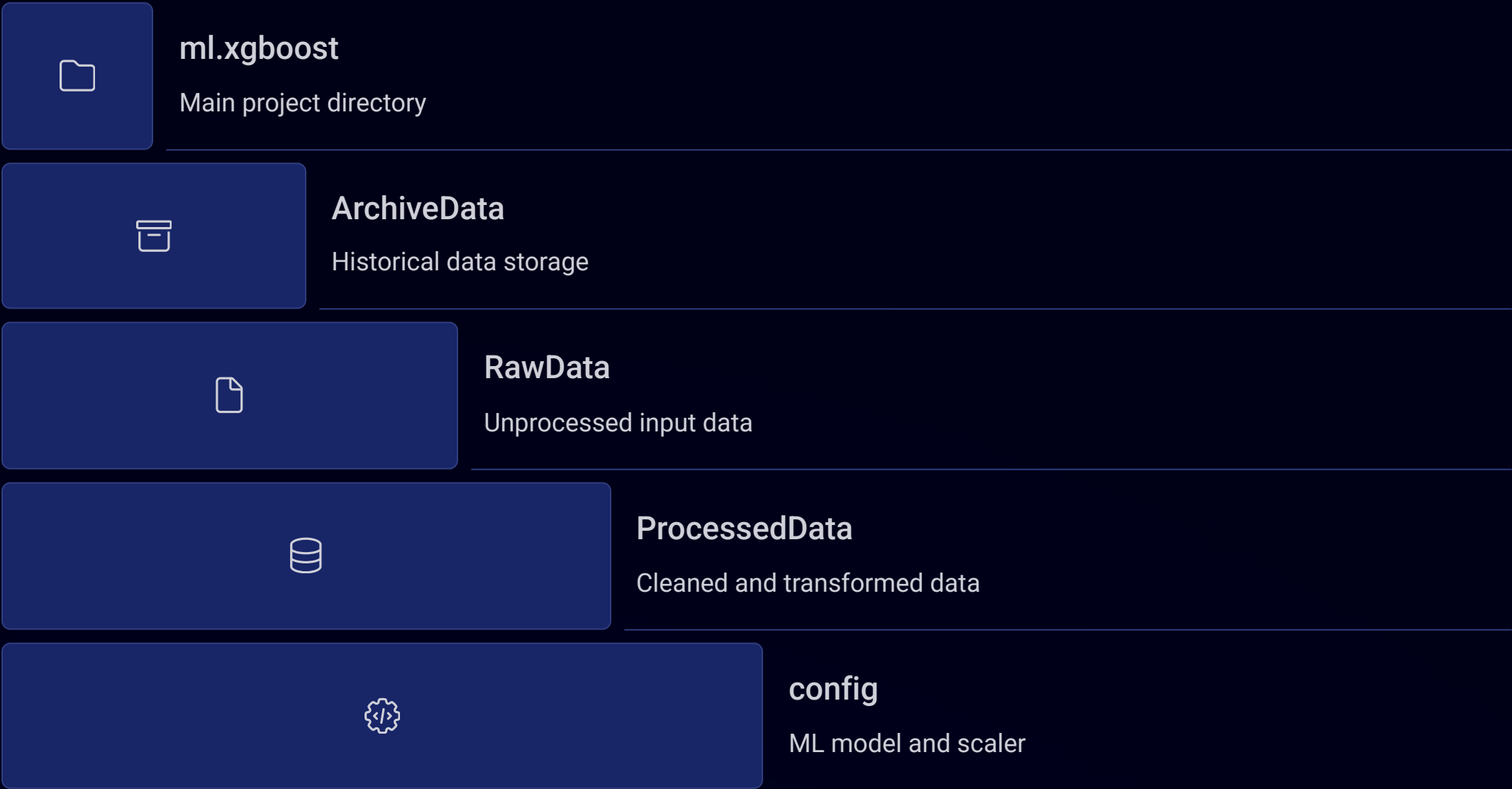
RMSE

Root Mean Square Error

€6.23

MAE

Mean Absolute Error

Model achieves excellent accuracy with low error rates. It has domain-specific features (lag, volatility, seasonality), applied grid search with time-based validation. Outputs include forecast plots, Excel exports, HTML dashboards.

S3 Bucket Configuration

ml.xgboost

Main project directory

ArchiveData

Historical data storage

RawData

Unprocessed input data

ProcessedData

Cleaned and transformed data

config

ML model and scaler

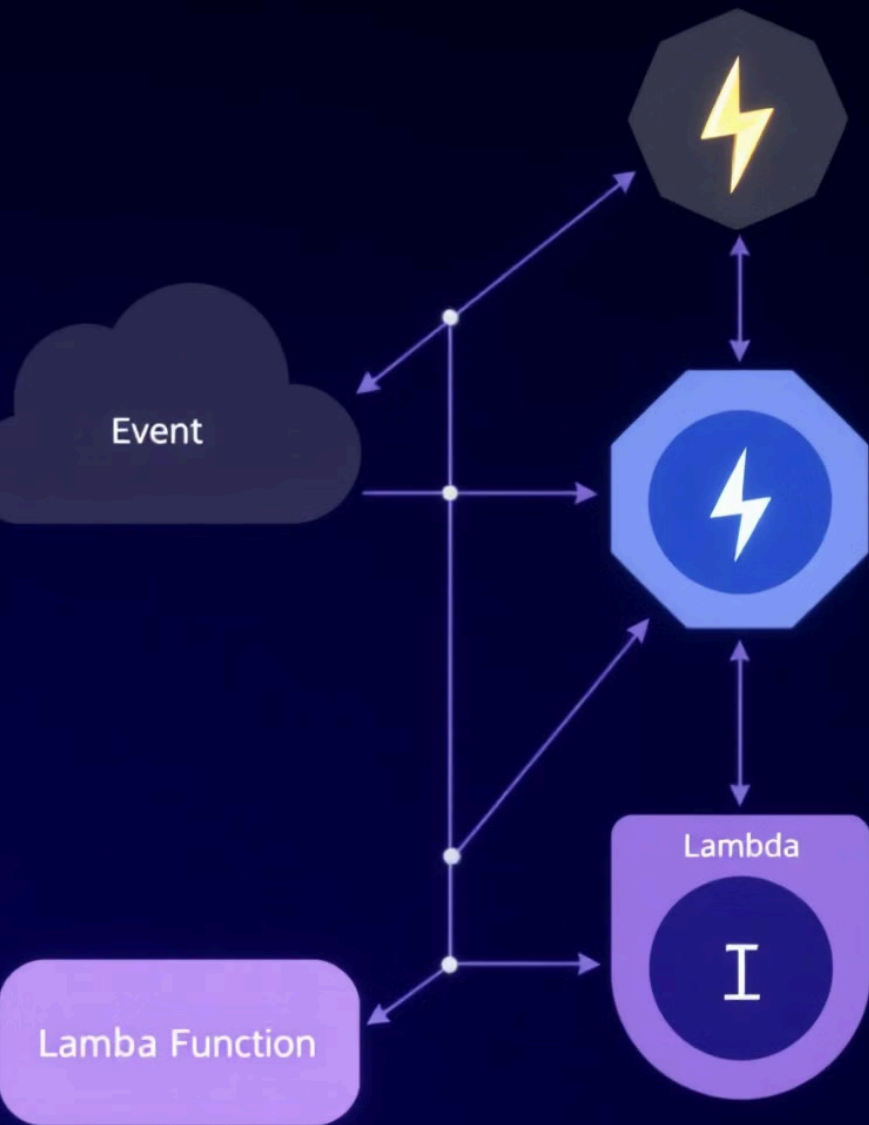# Lambda Functions

ml_xgboost_lambda

Data forecasting and analysis

ml_xgboost_email_lambda

Email delivery with SES

The system uses two specialized Lambda functions to handle forecasting and reporting tasks.

Automation Workflow

Sequence	Scope	Trigger	Type
1	Forecast execution	trigger_forecast_lambda	EventBridge/Scheduler
2	Send email with reports	trigger_email_lambda	S3 Event Notification

Tech Stack Overview

Core Components

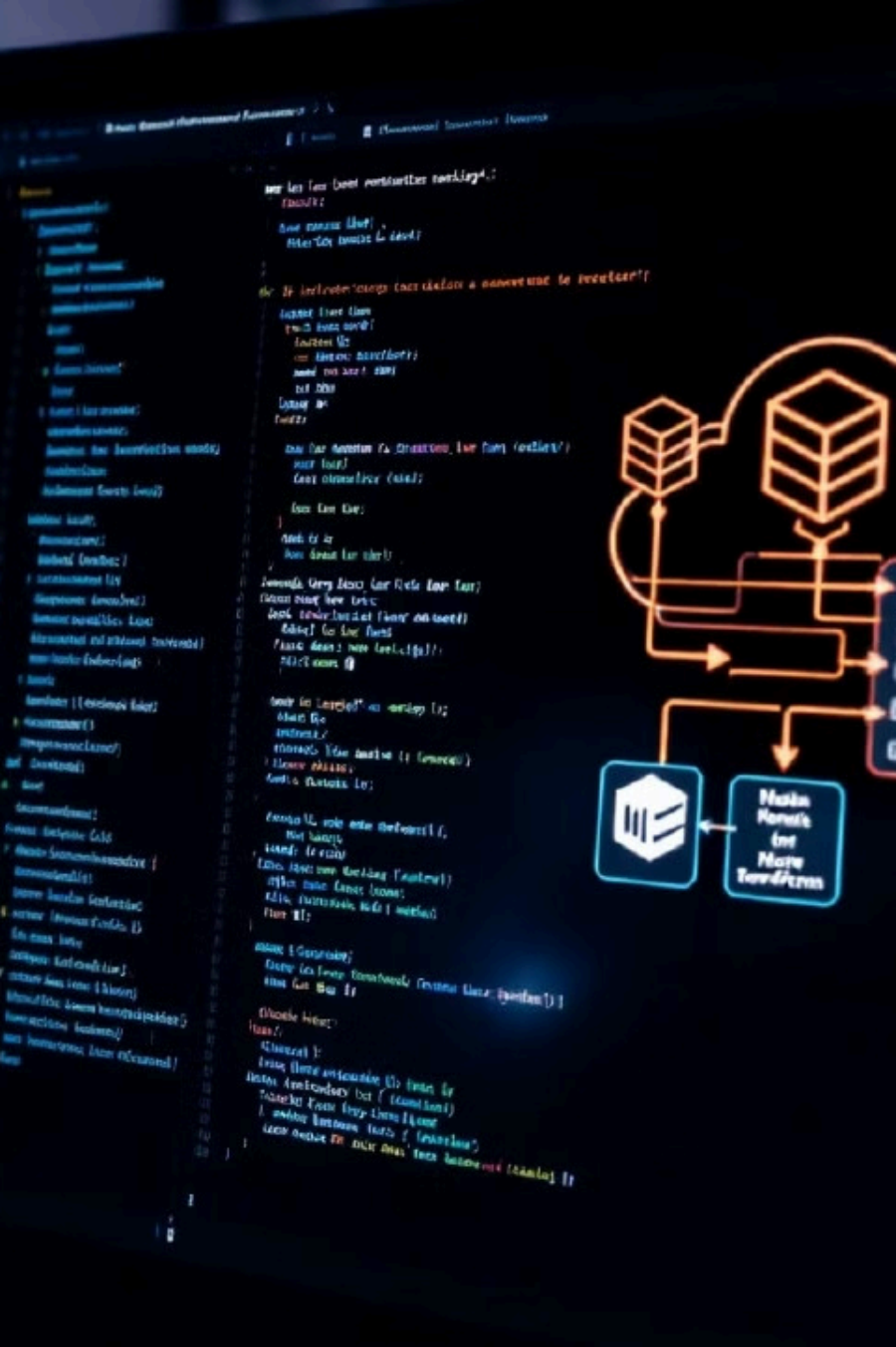- Python
- pandas, numpy
- xgboost
- sklearn, statsmodels

Visualization

- matplotlib
- plotly

AWS Services

- Lambda
- EventBridge
- S3
- SES

Deployment Architecture

Docker
Containerized deployment

AWS Lambda
Serverless execution

S3 Notifications
Event-based processing

EventBridge
Scheduled triggers

Infrastructure Management

Terraform

Infrastructure as code for consistent deployment.

AWS S3

Scalable storage for all data assets. Enables event-driven workflows.

AWS SES

Reliable email delivery for reports. Integrates seamlessly with Lambda functions.