Prediction of supermarket sales to reduce food waste

Cloud Computing I Presentation: Group 3 – 05.12.2025

esade

Agenda

The presentation is focused on 6 main parts

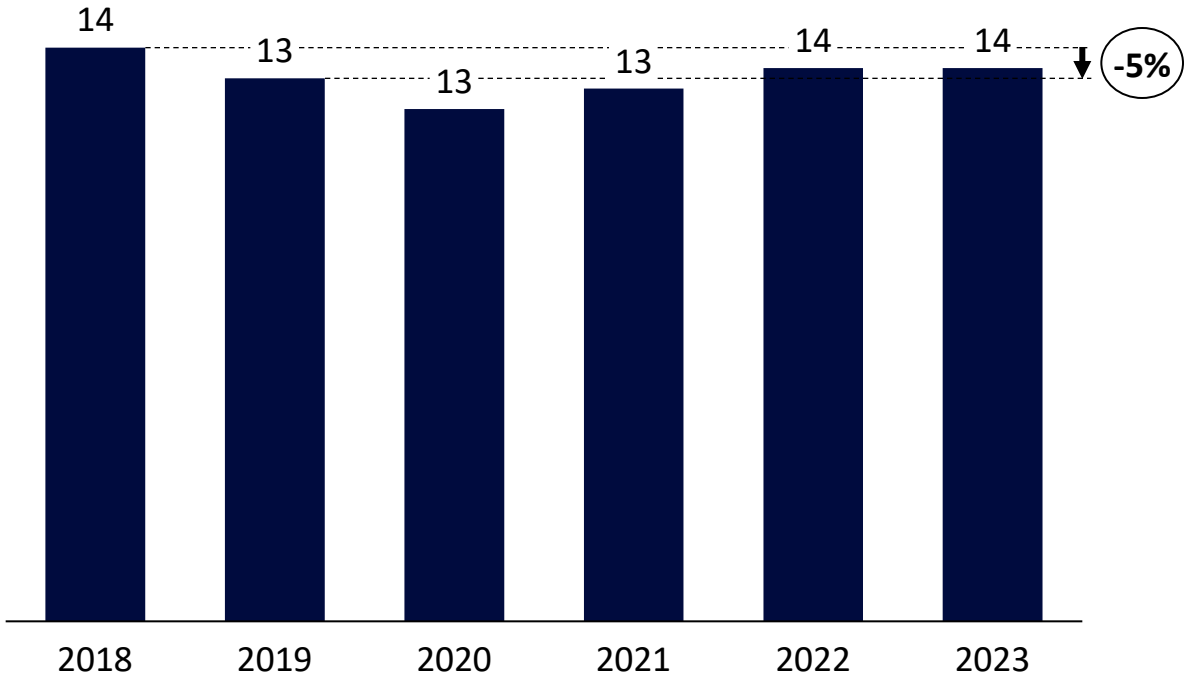
esade

1

Business need

Progress on reducing perishable good waste is slowing

Perishable food waste of UK retailers per year (in k tonnes)



Description

- Retailers are relying on good forecasts to buy the right amount of goods to avoid overstocking or selling out
- Good sales forecasts require well-trained models to buy just the right amount of goods
- These models need reliable data which includes aspects such as unique needs, new products, seasonal tastes, marketing etc.

The waste of perishable goods from supermarkets can be reduced by better forecasting the actual demand to avoid overstocking using an easily accessible ML model

Sources: Numan (2024)

2

Model explanation

Model was trained offline in 7 steps

esade

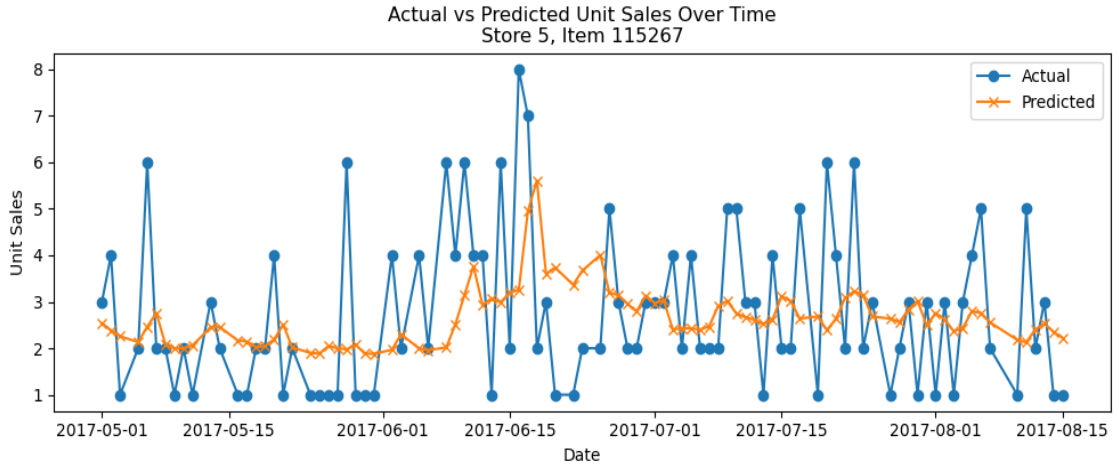
Most important steps of the model creation

1. **Subsampling & modelling tables:** Random sampling to handle size and merge of main sales data
2. **Time-based & lag feature engineering:** Transform raw daily sales into richer time-series feature set
3. **Target definition & split:** Define time-based unit sales and target and train/validation split
4. **Preprocessing pipeline with ColumnTransformer:** Transforming of numeric and categorical features
5. **Model zoo and hyperparameter grids:** Ridge & Lasso regression, RandomForest, XGB, LGBRegressor
6. **Model selection:** GridSearchCV searches over the hyperparameter grid
7. **Final model training and test prediction:** Refitting of best model and generation of predictions

Performance metrics

RMSE	MAE	R^2
9.89	3.45	0.74

Prediction of sales vs actual sales

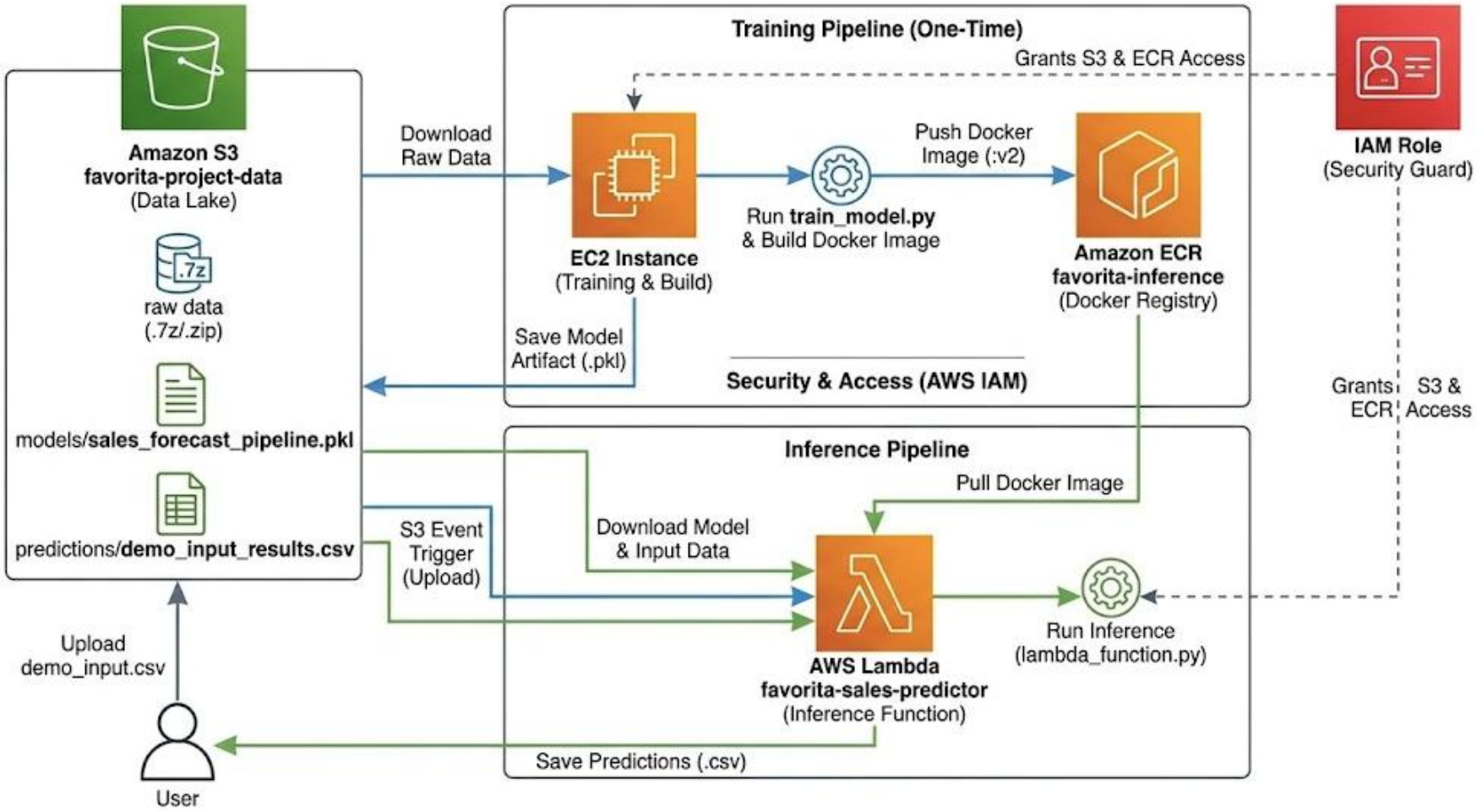


3

AWS Infrastructure

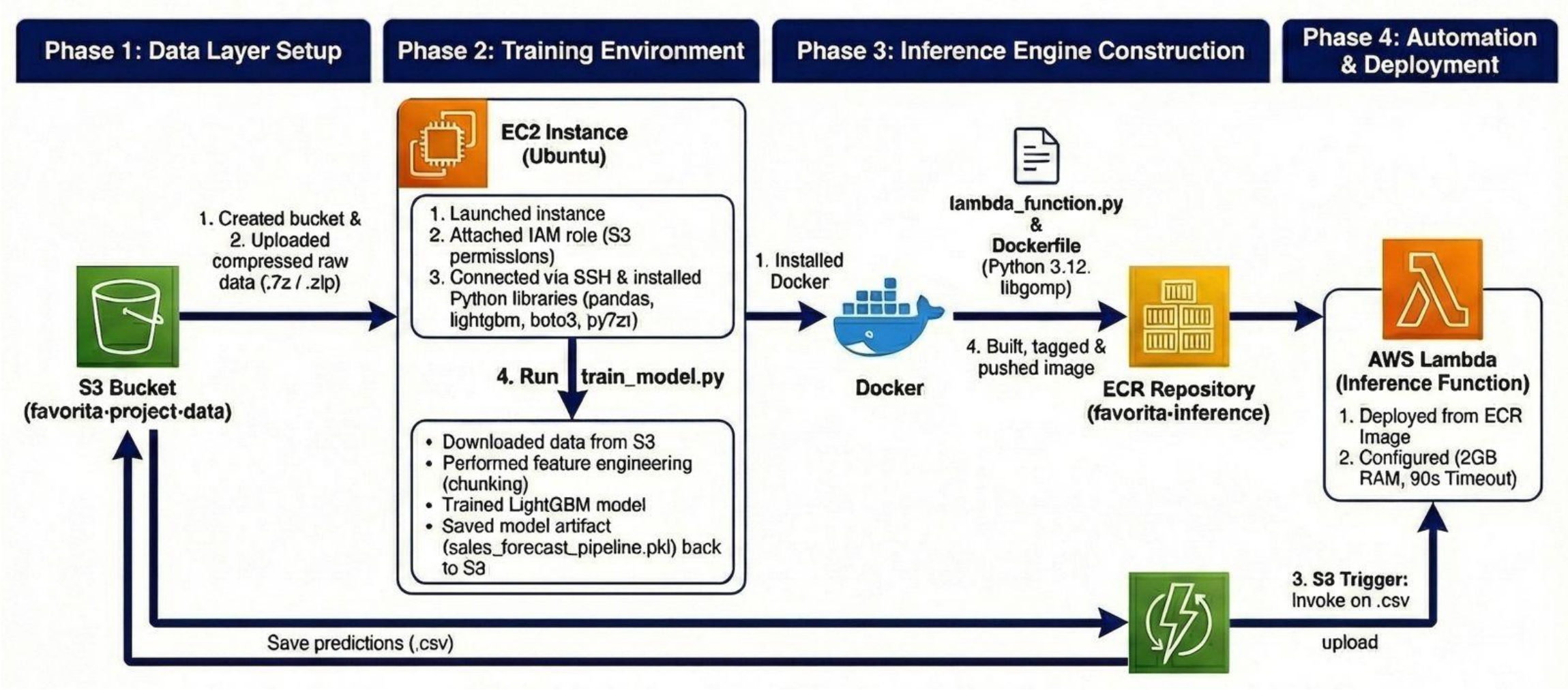
Overview of the AWS Infrastructure

esade



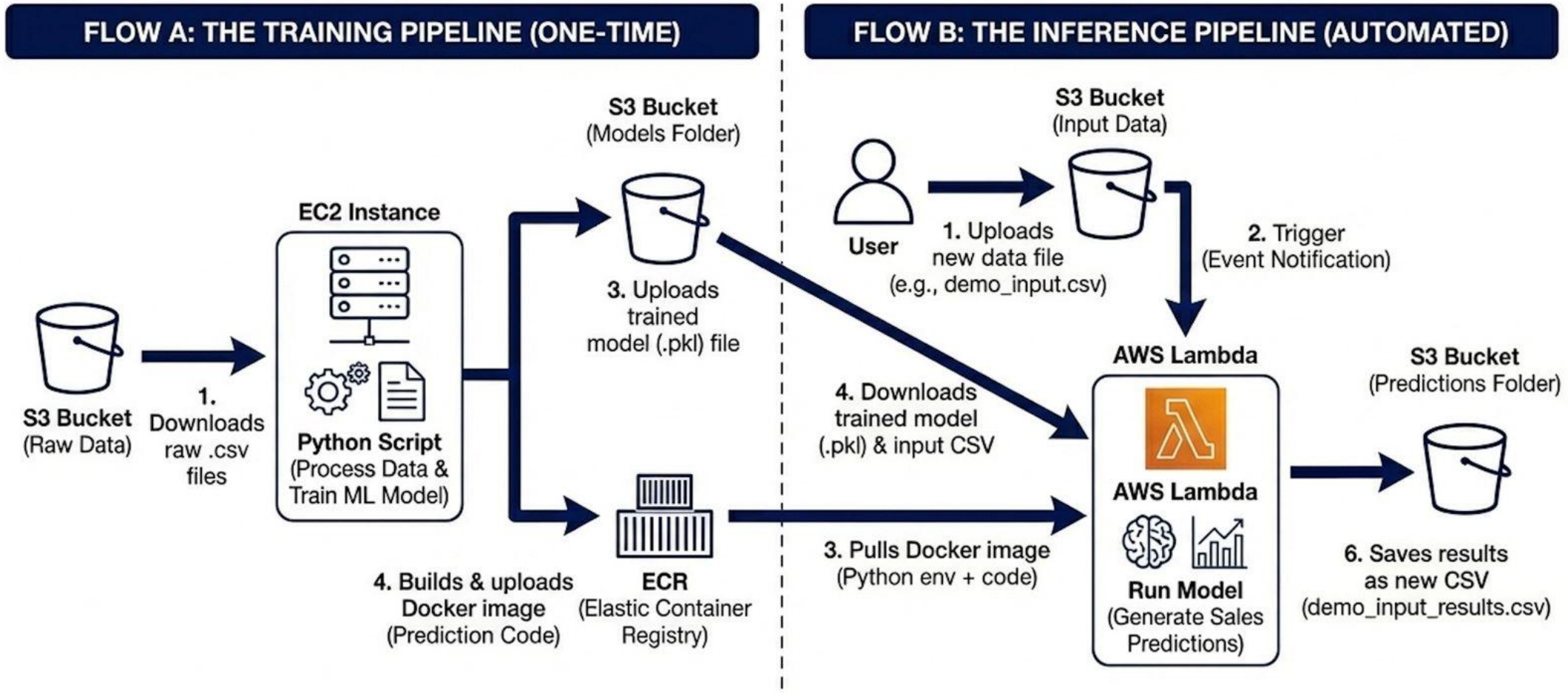
The development process can be divided into four phases

esade



Data flows between the different AWS components

esade



4

Live Demonstration

5

Key takeaways

The solution can be improved in the future

esade

1 Key takeaways

Cloud scalability: Project demonstrates the use of diverse AWS compute services to handle different workloads

2 Impact of solution

Reduction of food waste: By generating more accurate demand forecasts, managers avoid overstocking and therefore food waste

Operational efficiency: The “upload-and-forget” workflow removes the need for manual data processing

3 Improvement possibilities

Incorporation of external data: Adding of external signals such as local weather data, holiday calendars, or major local events

Add API gateway: Currently, the process is triggered by an S3 upload. Adding an API gateway would allow better access options

Inventory optimization layer: Move beyond raw prediction numbers by building a layer that translates sales into actual order quantities