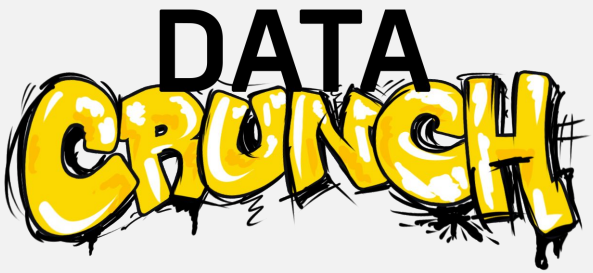
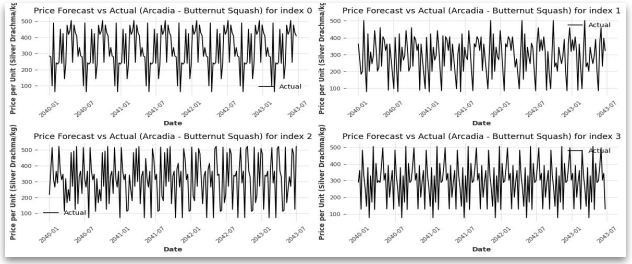


Data Crunch Final Round Presentation

Team ID	DataCrunch_045
Name of the Team	Helloworld 2.0
University, Faculty & Year	University of Moratuwa, Faculty of Engineering & Faculty of Information Technology, Third year
Team Members	Lasitha Amarasinghe
	Thakshaka Rathnayake
	Induwara Gayashan Morawakgoda

Problem Understanding & Data Analysis



- **Nature of the Data**
  - 25 Regions
  - 37 Commodities
  - 25 x 37 = 925 Combinations ⟶ 925 Time series
  - Each time series have different patterns
  - **Requires 925 individual models that are capable of retraining**

- **Pattern of the time series**
  - Duplicates data entries
  - Inconsistent frequency
  - Some weeks have more than one data point
  - Most data has a frequency of 6 or 7 days
  - **Requires consistency in time intervals**

DateDiff	
7.0	88848
6.0	86980
8.0	2436
5.0	2366
1.0	47

- **Impact of other features on Price**
  - Humidity, temperature, rainfall, and region contribute to a crop yield impact score
  - The crop yield impact score affects the price
  - **Requires incorporating weather impact into price prediction**

Model Selection & Feature Engineering

Model Selection

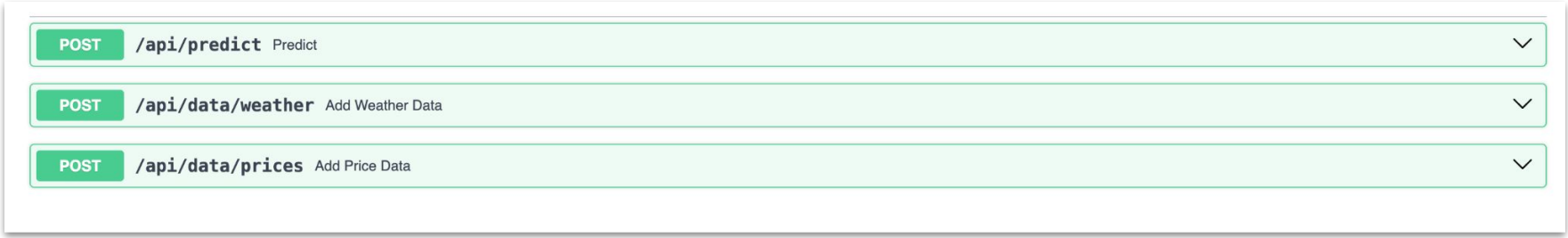
- **Chosen Model:** Darts N-BEATS Model.
- **Reasoning:** Have the strength in capturing long-term patterns in complex time series without requiring external covariates
- **Model Training:**Trained the model on historical data to predict the price for the next 4 weeks using a rolling window.

Feature Engineering

- Used weather data (temp, rainfall, humidity, region) to predict crop yield impact score.
- Merged with price data, cleaned, and resampled to weekly frequency.
- Used past 24 weeks of prices as input.
- Crop yield impact score used as a covariate in price prediction.

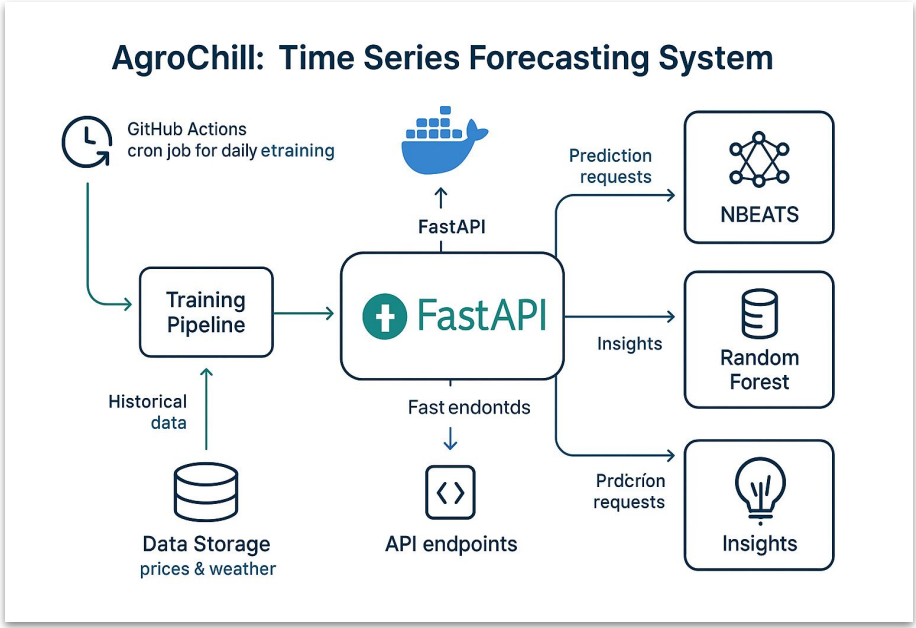
Data Pipeline Strategy

- **Automated Data Ingestion:** Weekly weather and price data are automatically pulled via APIs, ensuring continuous updates and storage in the datasets.



- **Periodic Re-Training:** The data pipeline adapts to new data by automatically retraining the models every week. GitHub Workflows are used to manage this process.

System Design



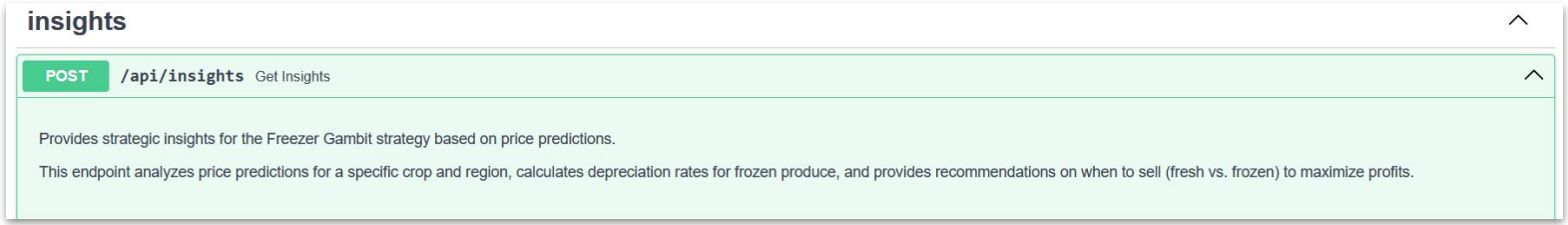
- **Architecture**
  - FastAPI Backend: Core API service providing predictions and insights, Scalable
  - Time Series Forecasting: N-BEATS model for price predictions up to 4 weeks ahead
  - Random Forest Model: Weather impact model to assess crop yield impact
  - Continuous Learning: Daily model retraining via GitHub Actions cron job
- **Data Flow**
  - Historical price and weather data → Training pipeline
  - New data ingestion via API endpoints
  - Prediction requests → Model inference → Strategic insights.
  - Automated retraining
- **Deployment**: Docker, GitHub Actions

Business Insights & Recommendations

- **Key Insights:**
  - Our system analyzes the trade-off between immediate sales and strategic freezing
  - Calculates compound depreciation rates for frozen produce (10% weekly)
  - Identifies optimal selling windows across 4-week forecasts

- **Strategic Recommendations:**
  - Freeze crops when future fresh prices exceed depreciated frozen values
  - Sell immediately when depreciation outpaces market growth
  - Target high-volatility markets for maximum freezing benefits

- **Business Impact:**
  - Increases profit margins through strategic timing
  - Reduces waste by extending produce shelf life
  - Optimizes cold storage utilization based on data-driven decisions
  - Provides competitive advantage through predictive market intelligence



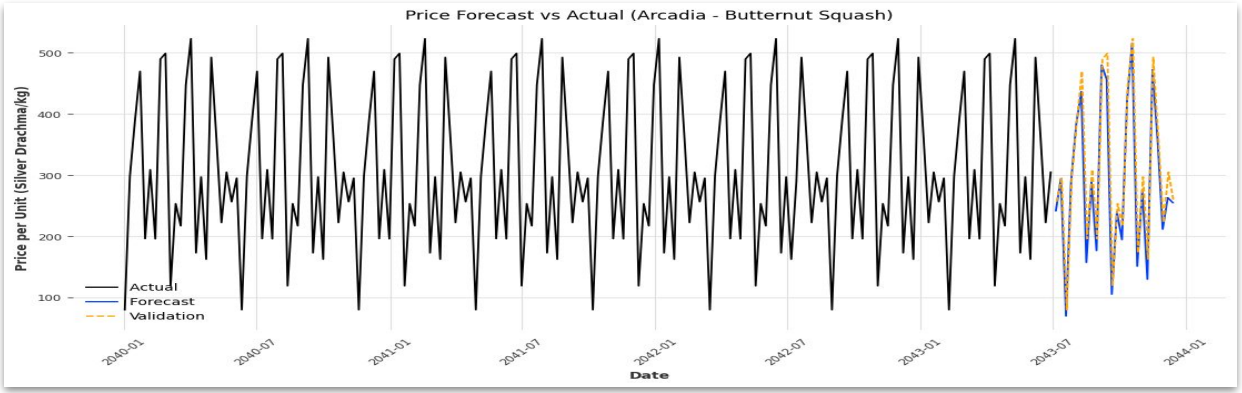
```
{
  "crop": "Butternut Squash",
  "region": "Gotham",
  "crop_type": "Vegetable",
  "price_projections": [
    {
      "week": 1,
      "date": "2043-07-07",
      "fresh_price": 168,
      "frozen_price": 53.51,
      "depreciation_percentage": 10
    },
    {
      "week": 2,
      "date": "2043-07-14",
      "fresh_price": 260.49,
      "frozen_price": 48.16,
      "depreciation_percentage": 19
    },
    {
      "week": 3,
      "date": "2043-07-21",
      "fresh_price": 362,
      "frozen_price": 43.35,
      "depreciation_percentage": 27.1
    },
    {
```

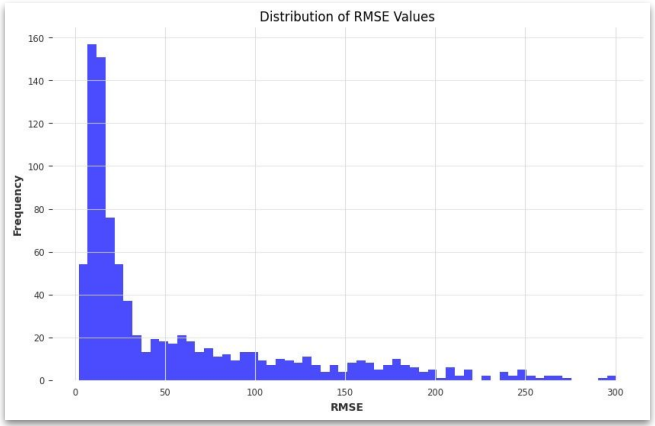

Download


```
    },
    {
      "week": 3,
      "date": "2043-07-21",
      "fresh_price": 362,
      "frozen_price": 43.35,
      "depreciation_percentage": 27.1
    },
    {
      "week": 4,
      "date": "2043-07-28",
      "fresh_price": 271.04,
      "frozen_price": 39.01,
      "depreciation_percentage": 34.39
    }
  ],
  "best_selling_week": 3,
  "best_selling_date": "2043-07-21",
  "best_selling_price": 362,
  "is_freezing_recommended": true,
  "recommendation": "Freeze the Butternut Squash and sell in week 3 (2043-07-21) for 362.0 Silver Drachma/kg.",
  "insights": [
    "Prices for Butternut Squash in Gotham show an upward trend over the next 4 weeks.",
    "Freezing will yield approximately 735.06% more profit compared to selling frozen produce.",
    "The market for Butternut Squash in Gotham shows high volatility, suggesting careful timing of sales."
  ]
}
```


Download

Model Performance





count	925.000000
mean	55.981418
std	62.949578
min	2.086860
25%	12.616102
50%	24.271516
75%	80.824895
max	300.311302

Thank You!