QUANTRA
DATA ANALYSTS
MANUFACTURING DOWDTIME

Manufacturing Downtime Analysis

Mining the Past, Powering the Future.

Digital Egypt Pioneers

وزارة الاتصـــــالات
وتكنولوجيا المعلومات

Data Analyst Specalist
Track

Under Supervision of

CLS Learning Solutions

Instructor: Dr. Maged Magdy

Manufacturing Downtime

Mining the Past, Powering the Future.

QUANTRA
DATA ANALYSTS

QUANTRA
DATA ANALYSTS
MANUFACTURING DOWDTIME

Our Team

Hoda Elkassas

Hanaa Edreis

Farah Sameh

Yomna Ayman

Rania Elwan

QUANTRA
DATA ANALYSTS
MANUFACTURING DOWDTIME

Introduction

Problem Statment

Manufacturing performance is heavily impacted by unplanned downtime, which causes noticeable drops in productivity and efficiency. Many factories still lack clear visibility into the real reasons behind these interruptions, making it difficult to identify high-risk operators, machines, or recurring patterns.

Without a reliable way to forecast downtime for the next day, production teams face challenges in scheduling, resource allocation, and maintenance planning. This leads to delayed orders, higher operational costs, and reduced overall performance.

Our project highlights the importance of understanding downtime behavior and building a predictive system that supports smarter, data-driven decision-making.

Objectives

- Analyze historical downtime patterns and identify the most influential contributing factors.
- Measure downtime behavior across operators, product sizes, flavors, and shifts to highlight high-risk areas.
- Build a machine learning regression model to predict downtime minutes per production batch.
- Generate a reliable next-day downtime forecast using engineered time-based and operational features.
- Provide managers with clear, actionable insights** through an interactive dashboard (factors, operators, shift risks).
- Deploy predictions via a REST API to enable smooth integration between the model and the dashboard.

QUANTRA
DATA ANALYSTS
MANUFACTURING DOWDTIME

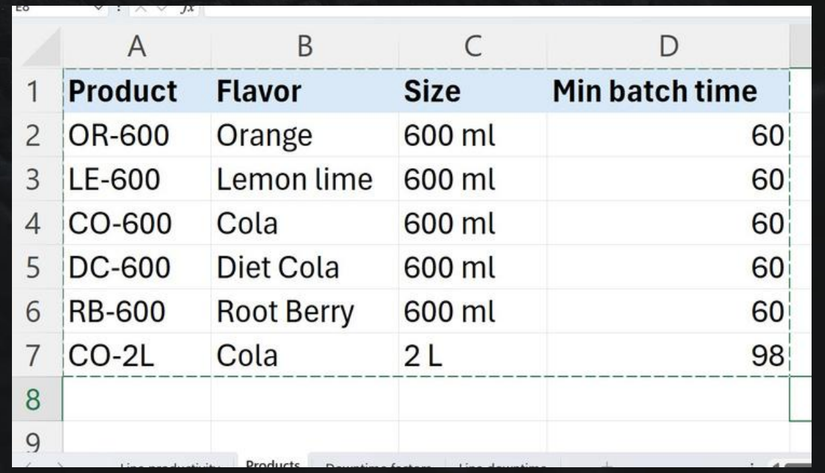
Scope

- Utilize historical factory datasets, including batch logs, product specifications, operators, and downtime factors.
- Perform comprehensive data cleaning (invalid values, missing data, size standardization, operator normalization).
- Apply feature engineering (downtime factor encoding, time features, aggregated batch metrics).
- Develop and evaluate multiple ML models (Random Forest, XGBoost) using MAE, RMSE, and R^2.
- Build a structured deployment pipeline: Model → API Endpoint (/predict) → Interactive Dashboard.
- Visualize predictions and insights through a central dashboard covering risk metrics, factor breakdown, and forecasted downtime.

Dataset Overview

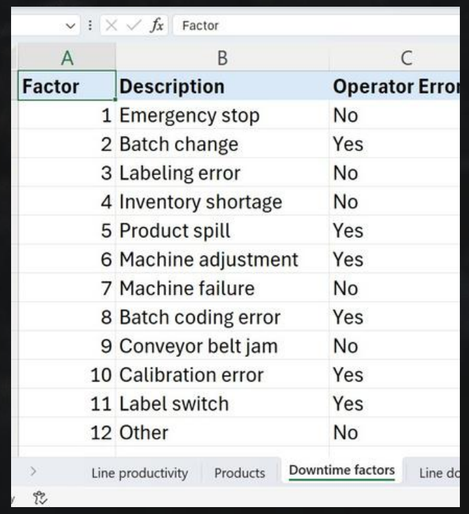
Data Volume

We worked with 38 production batch records collected from real manufacturing operations.
The data captures product details, operator performance, batch timings, and downtime distributions across multiple factors.



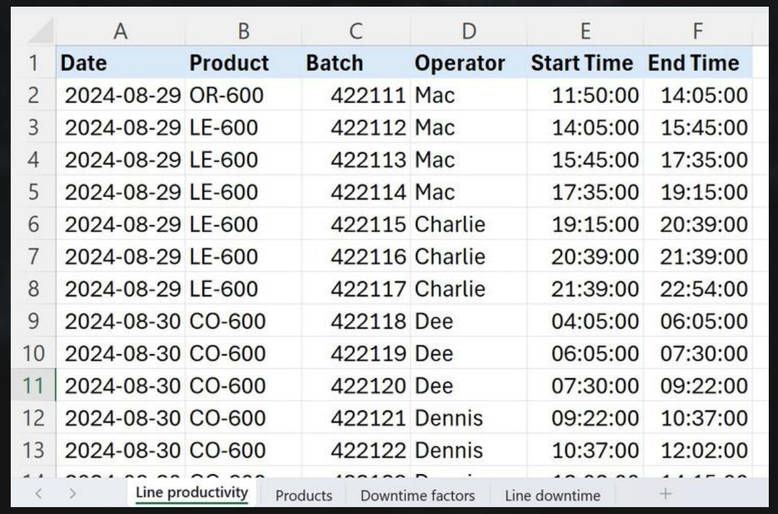
	A	B	C	D
1	Product	Flavor	Size	Min batch time
2	OR-600	Orange	600 ml	60
3	LE-600	Lemon lime	600 ml	60
4	CO-600	Cola	600 ml	60
5	DC-600	Diet Cola	600 ml	60
6	RB-600	Root Berry	600 ml	60
7	CO-2L	Cola	2 L	98
8				
9				

Products



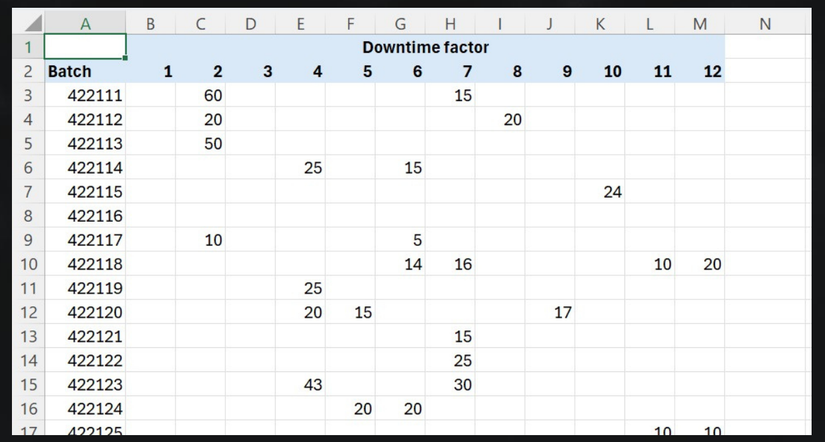
	A	B	C
	Factor	Description	Operator Error
1	1	Emergency stop	No
2	2	Batch change	Yes
3	3	Labeling error	No
4	4	Inventory shortage	No
5	5	Product spill	Yes
6	6	Machine adjustment	Yes
7	7	Machine failure	No
8	8	Batch coding error	Yes
9	9	Conveyor belt jam	No
10	10	Calibration error	Yes
11	11	Label switch	Yes
12	12	Other	No

Line productivity Products Downtime factors Line do



	A	B	C	D	E	F
1	Date	Product	Batch	Operator	Start Time	End Time
2	2024-08-29	OR-600	422111	Mac	11:50:00	14:05:00
3	2024-08-29	LE-600	422112	Mac	14:05:00	15:45:00
4	2024-08-29	LE-600	422113	Mac	15:45:00	17:35:00
5	2024-08-29	LE-600	422114	Mac	17:35:00	19:15:00
6	2024-08-29	LE-600	422115	Charlie	19:15:00	20:39:00
7	2024-08-29	LE-600	422116	Charlie	20:39:00	21:39:00
8	2024-08-29	LE-600	422117	Charlie	21:39:00	22:54:00
9	2024-08-30	CO-600	422118	Dee	04:05:00	06:05:00
10	2024-08-30	CO-600	422119	Dee	06:05:00	07:30:00
11	2024-08-30	CO-600	422120	Dee	07:30:00	09:22:00
12	2024-08-30	CO-600	422121	Dennis	09:22:00	10:37:00
13	2024-08-30	CO-600	422122	Dennis	10:37:00	12:02:00

Line productivity Products Downtime factors Line downtime



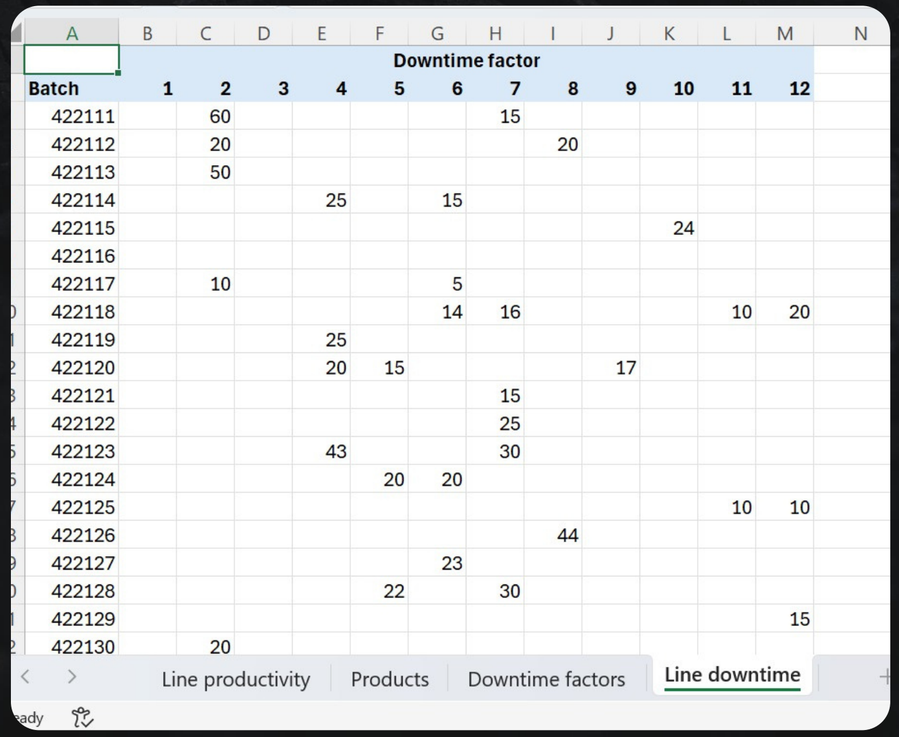
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1		Downtime factor												
2	Batch	1	2	3	4	5	6	7	8	9	10	11	12	
3	422111		60					15						
4	422112		20						20					
5	422113		50											
6	422114				25		15							
7	422115										24			
8	422116													
9	422117		10				5							
10	422118						14	16				10	20	
11	422119				25									
12	422120				20	15				17				
13	422121							15						
14	422122							25						
15	422123				43			30						
16	422124					20	20							
17	422125											10	10	

QUANTRA
DATA ANALYSTS
MANUFACTURING DOWNTIME

Extracting temporal features (Start Hour, Day of Week) and operator risk metrics (Average Downtime per Operator).

Data Preprocessing

Before Handleing & Cleaning



		Downtime factor											
	Batch	1	2	3	4	5	6	7	8	9	10	11	12
	422111		60					15					
	422112		20						20				
	422113		50										
	422114				25		15						
	422115										24		
	422116												
	422117		10				5						
0	422118						14	16				10	20
1	422119				25								
2	422120				20	15				17			
3	422121							15					
4	422122							25					
5	422123				43			30					
6	422124					20	20						
7	422125											10	10
8	422126								44				
9	422127						23						
0	422128					22		30					
1	422129												15
2	422130		20										

Line productivity | Products | Downtime factors | Line downtime

Ready

After Handleing & Cleaning

	A	B	C
1	Batch	Factor	Downtime
2	422111	2	60
3	422112	2	20
4	422113	2	50
5	422117	2	10
6	422130	2	20
7	422138	3	20
8	422145	3	22

Data Preprocessing

Before Handleing & Cleaning

18	2024-08-30	CO-600	422127	Charlie	18:59:00	20:22:00
19	2024-08-30	CO-600	422128	Charlie	20:22:00	22:14:00
20	2024-08-30	CO-600	422129	Charlie	22:14:00	23:29:00
21	2024-08-31	CO-600	422130	Dee	07:45:00	09:05:00
22	2024-08-31	CO-600	422131	Dee	09:05:00	10:35:00
23	2024-08-31	CO-600	422132	Dee	10:35:00	11:35:00
24	2024-08-31	DC-600	422133	Dee	11:35:00	12:55:00
25	2024-08-31	DC-600	422134	Mac	12:55:00	14:45:00
26	2024-08-31	DC-600	422135	Mac	14:45:00	16:30:00
27	2024-08-31	DC-600	422136	Mac	16:30:00	17:30:00
28	2024-09-02	RB-600	422137	Dee	01:00:00	02:45:00
29	2024-09-02	RB-600	422138	Dee	02:45:00	04:05:00
30	2024-09-02	RB-600	422139	Dee	04:05:00	05:40:00
31	2024-09-02	RB-600	422140	Dee	05:40:00	07:43:00
32	2024-09-02	RB-600	422141	Dennis	07:43:00	08:50:00
33	2024-09-02	RB-600	422142	Dennis	08:50:00	10:20:00
34	2024-09-02	RB-600	422143	Dennis	10:20:00	12:18:00
35	2024-09-02	CO-2L	422144	Dennis	12:18:00	14:50:00
36	2024-09-02	CO-2L	422145	Charlie	14:50:00	16:50:00
37	2024-09-02	CO-2L	422146	Charlie	16:50:00	19:30:00
38	2024-09-02	CO-2L	422147	Charlie	19:30:00	22:55:00
39	2024-09-03	CO-2L	422148	Mac	22:55:00	01:05:00

Line productivity | Products | Downtime factors | Line downtime

After Handleing & Cleaning

End Date	Production time Minutes	Start_timestamp	End_timestamp
Friday, August 30, 2024	112	8/30/2024 8:22:00 PM	8/30/2024 10:14:00 PM
Friday, August 30, 2024	75	8/30/2024 10:14:00 PM	8/30/2024 11:29:00 PM
Saturday, August 31, 2024	80	8/31/2024 7:45:00 AM	8/31/2024 9:05:00 AM
Saturday, August 31, 2024	90	8/31/2024 9:05:00 AM	8/31/2024 10:35:00 AM
Saturday, August 31, 2024	60	8/31/2024 10:35:00 AM	8/31/2024 11:35:00 AM
Saturday, August 31, 2024	80	8/31/2024 11:35:00 AM	8/31/2024 12:55:00 PM
Saturday, August 31, 2024	110	8/31/2024 12:55:00 PM	8/31/2024 2:45:00 PM
Saturday, August 31, 2024	105	8/31/2024 2:45:00 PM	8/31/2024 4:30:00 PM
Saturday, August 31, 2024	60	8/31/2024 4:30:00 PM	8/31/2024 5:30:00 PM
onday, September 2, 2024	105	9/2/2024 1:00:00 AM	9/2/2024 2:45:00 AM
onday, September 2, 2024	80	9/2/2024 2:45:00 AM	9/2/2024 4:05:00 AM
onday, September 2, 2024	95	9/2/2024 4:05:00 AM	9/2/2024 5:40:00 AM
onday, September 2, 2024	123	9/2/2024 5:40:00 AM	9/2/2024 7:43:00 AM
onday, September 2, 2024	67	9/2/2024 7:43:00 AM	9/2/2024 8:50:00 AM
onday, September 2, 2024	90	9/2/2024 8:50:00 AM	9/2/2024 10:20:00 AM
onday, September 2, 2024	118	9/2/2024 10:20:00 AM	9/2/2024 12:18:00 PM
onday, September 2, 2024	152	9/2/2024 12:18:00 PM	9/2/2024 2:50:00 PM
onday, September 2, 2024	120	9/2/2024 2:50:00 PM	9/2/2024 4:50:00 PM
onday, September 2, 2024	160	9/2/2024 4:50:00 PM	9/2/2024 7:30:00 PM
onday, September 2, 2024	205	9/2/2024 7:30:00 PM	9/2/2024 10:55:00 PM
esday, September 4, 2024	130	9/3/2024 10:55:00 PM	9/4/2024 1:05:00 AM

QUANTRA
DATA ANALYSTS
MANUFACTURING DOWNTIME

Data Preprocessing

Before Handleing & Cleaning

	Downtime factor											
Batch	1	2	3	4	5	6	7	8	9	10	11	12
422111		60					15					
422112		20						20				
422113		50										
422114				25		15						
422115										24		
422116												
422117		10				5						
422118						14	16				10	20
422119				25								
422120				20	15				17			
422121							15					
422122							25					
422123				43			30					
422124					20	20						
422125											10	10
422126								44				
422127						23						
422128					22		30					
422129												15
422130		20										
422131				20						10		

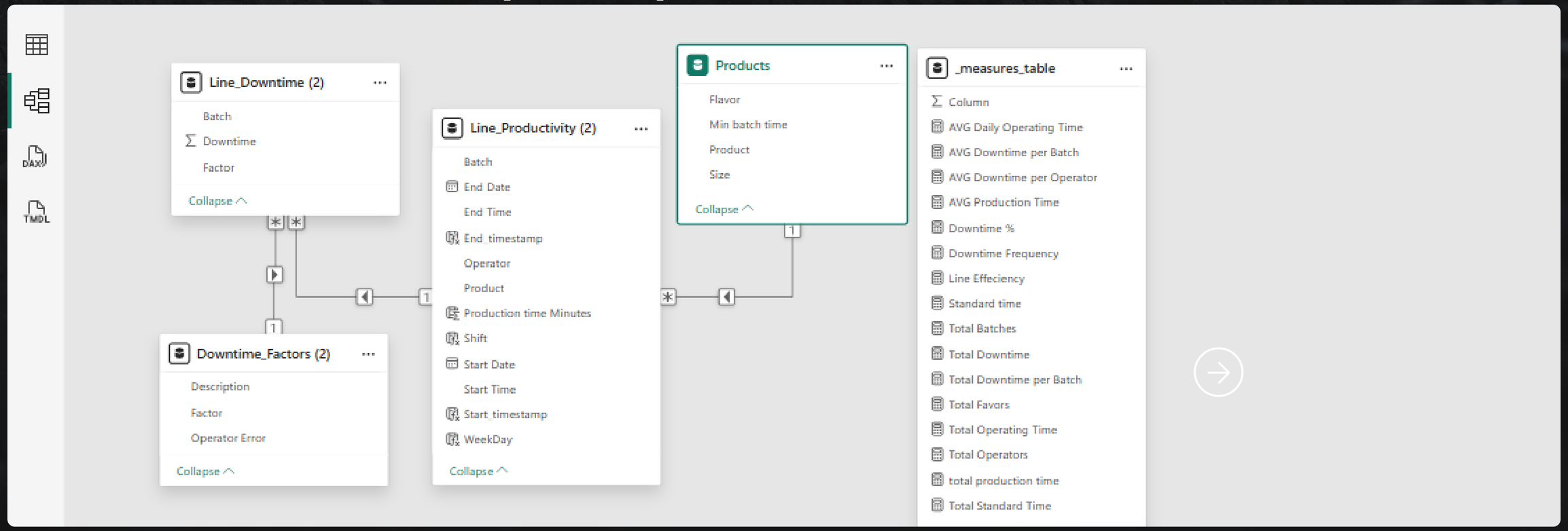
After Handleing & Cleaning

422130	2	20
422131	4	20
422131	10	10
422133	7	20
422134	7	30
422134	8	20
422135	4	30
422135	12	15

Data Preprocessing

Extracting temporal features (Start Hour, Day of Week) and operator risk metrics (Average Downtime per Operator).

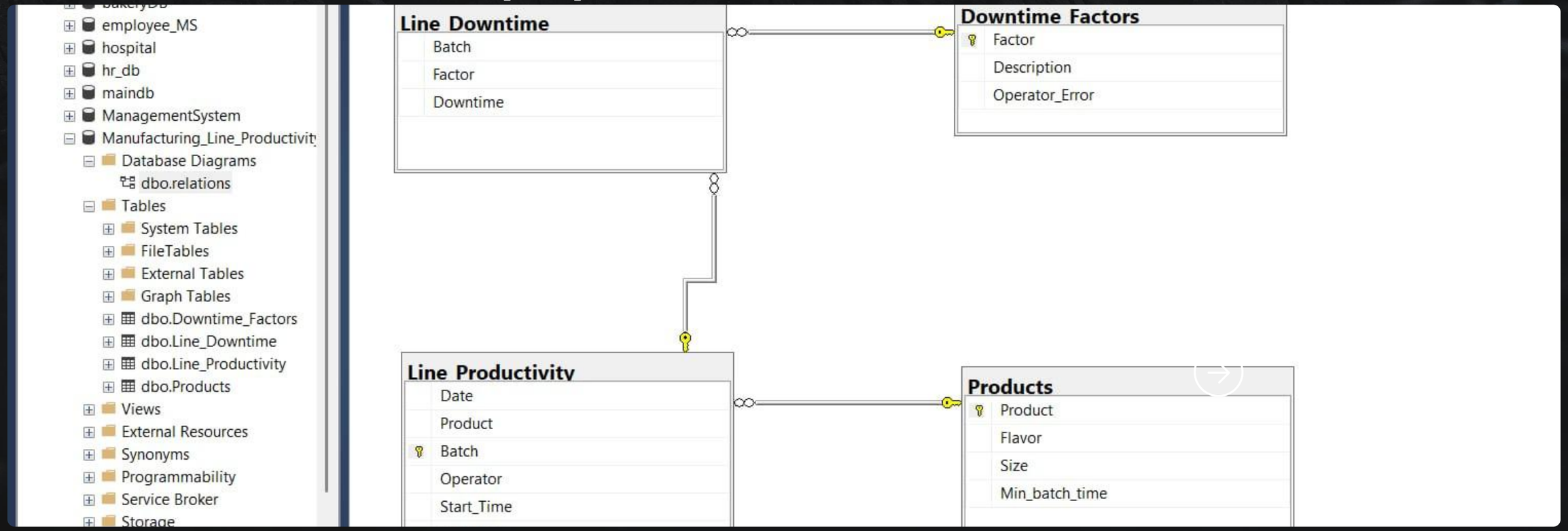
Relations betwwen Tables (Power BI)



Data Modeling

Extracting temporal features (Start Hour, Day of Week) and operator risk metrics (Average Downtime per Operator).

Relations betwwen Tables (SQL)



QUANTRA
DATA ANALYSTS
MANUFACTURING DOWNTIME

Exploratory Data Analysis-EDA

Operator Distribution

- Identified the most active operators and highlighted those contributing the highest downtime.

Downtime Factors

- A few factors cause most of the downtime. Mechanical issues, material jams, and changeovers appear most frequently.

Product Analysis

- Larger sizes (600 ml) show higher batch time and downtime. Most produced products also affect total line load.

Temporal Analysis

- Downtime varies across days and shifts. Clear peak days with high downtime and others with higher productivity.

Line Utilization

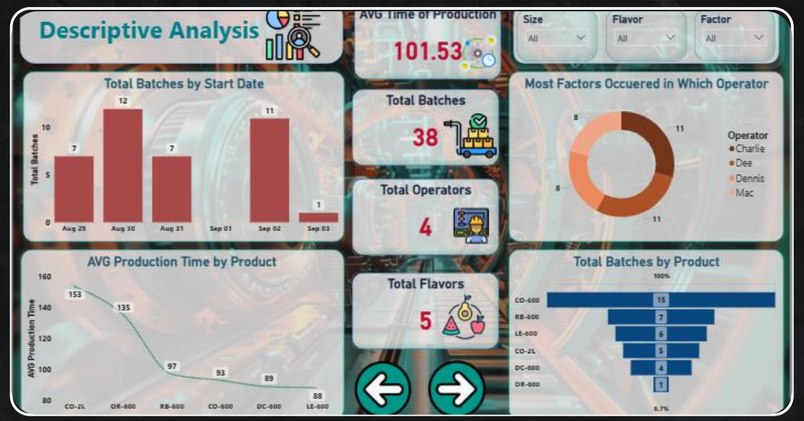
- Daily runtime fluctuates, and downtime represents a noticeable portion of total production time.

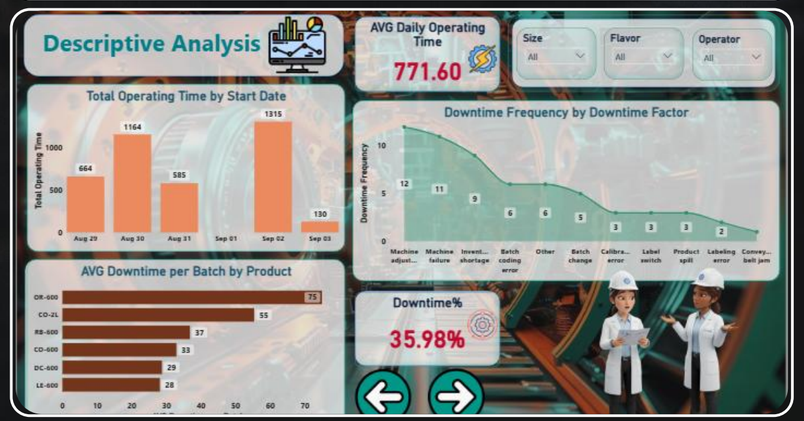
QUANTRA
DATA ANALYSTS
MANUFACTURING DOWDTIME

Visualization

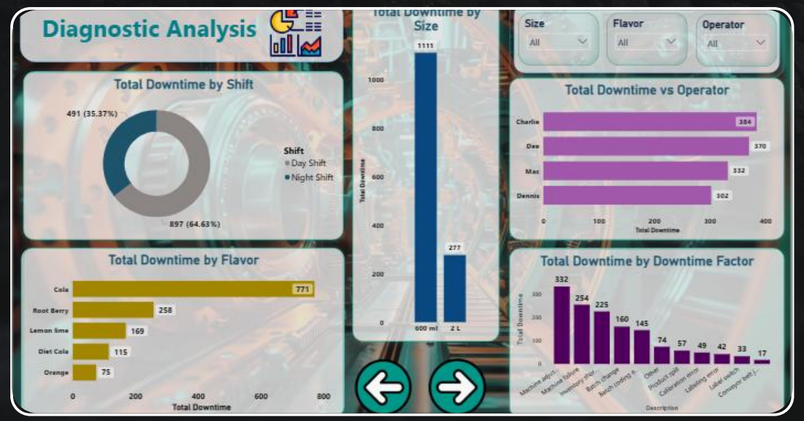
To see the Analysis clickhere

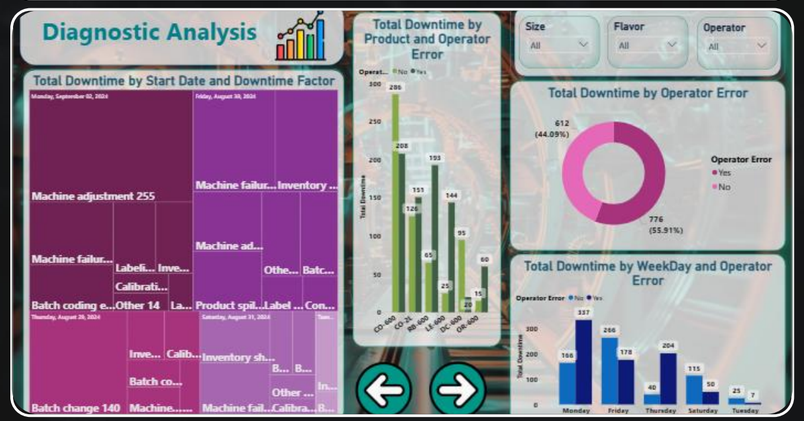
1 Discriptive Analysis



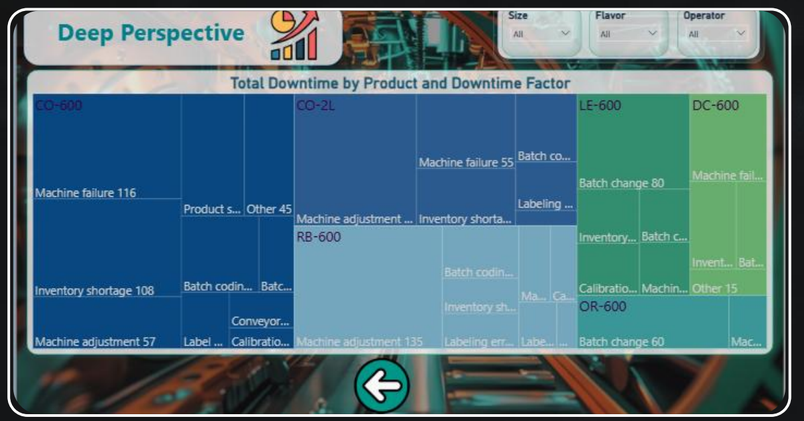


2 Diagnostic Analysis





3 Deep Perspective



AI Chatbot

Overview

Our AI chatbot connects directly to our operational data, allowing users to ask natural-language questions and instantly receive clear, data-driven answers supported by autogenerated charts and insights.

Key Features

- Conversational Data Access: Users can query production, downtime, or performance metrics through simple chat prompts.
- Automated Visualizations: The chatbot generates tailored charts to illustrate trends, comparisons, or anomalies.
- Integrated Dashboard: A unified interface displays key KPIs, recommended actions, and historical insights.
- Smart Insights: The system highlights patterns and provides context behind results.

QUANTRA
DATA ANALYSTS
MANUFACTURING DOWDTIME

AI Chatbot

Benefits

- Faster Decision-Making: Eliminates manual data searching and speeds up access to critical insights.

Adaptive & Scalable: Learns from user interactions and can expand with future data sources

- Improved Accuracy: Reduces human interpretation errors by using consistent, up-to-date data sources.

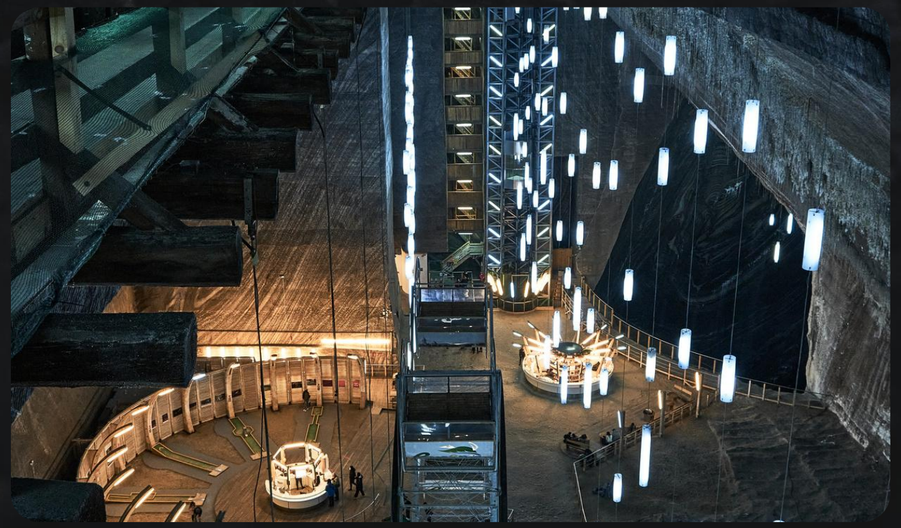
- User-Friendly: Accessible for both technical and non-technical staff, promoting wider data adoption.

Adaptive & Scalable: Learns from user interactions and can expand with future data sources or models.

AI Chatbot


Machine adjustment
Machine failure
Inventory shortage
Batch change
Batch coding error
Other
Product add
Calibration error
Labeling error
Label switch
Cause
0 50 100 150 200 250 300
Total Downtime (min)
Charlie
Dee
Mac
Dennis
Operator
0 50 100 150 200 250 300
Total Downtime (min)
Efficiency by Product
Downtime Over Time
OR-500
LE-500
CO-500
500
400

Data Preprocessing for ML



Data Merging

Three primary tables were joined using Batch ID and Product ID.

Datetime Handling

Correcting "Overnight Batches" (where End Time < Start Time) and calculating the actual Batch Duration in minutes.

Target Aggregation

Three primary tables were joined using Batch ID and Product ID.

Feature Engineering

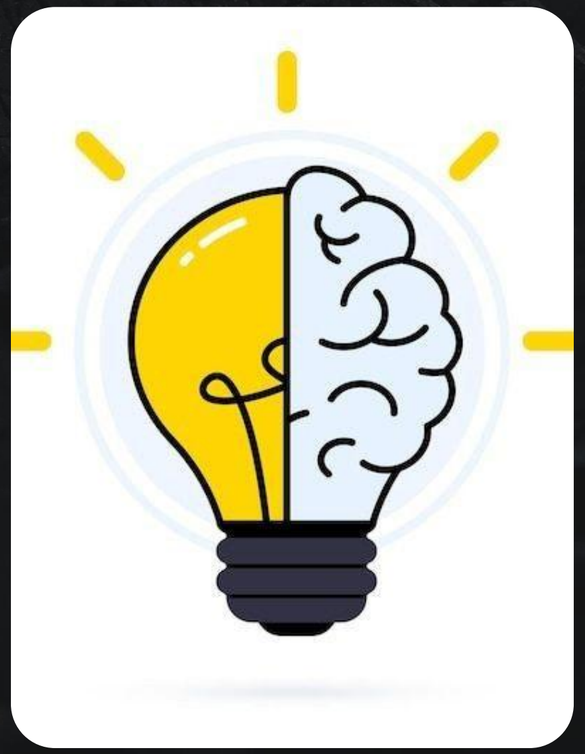
Extracting temporal features (Start Hour, Day of Week) and operator risk metrics (Average Downtime per Operator).

Statistical Treatment

Correcting data formats and handling Missing Values (by filling downtime factors with zero).

QUANTRA
DATA ANALYSTS
MANUFACTURING DOWDTIME

Machine Learning Model

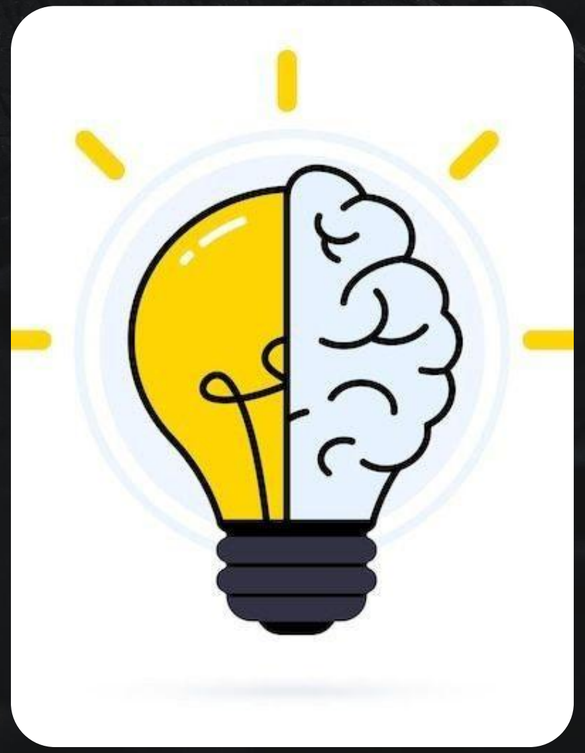


Overview:

The Machine Learning module predicts Duration Deviation — how many minutes the next batch is expected to be delayed.
It uses advanced engineered features that capture operator behavior, downtime patterns, product complexity, and time-based performance signals.

QUANTRA
DATA ANALYSTS
MANUFACTURING DOWDTIME

Machine Learning Model

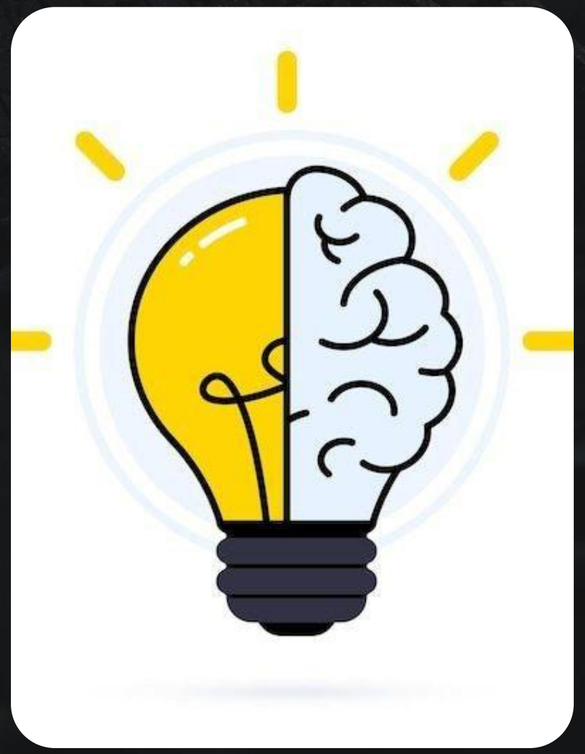


Key Features:

- Predictive modeling using a Random Forest Regressor
- Uses 10+ engineered features such as:
  - Lagged downtime history
  - Operator-product interaction behaviors
  - Micro-stops and incident frequency
  - High-risk failure indicators
  - Operator fatigue sequence
- Encodes categorical variables to numerical vectors
- Evaluated using MAE, R^2, and 5-Fold Cross-Validation

QUANTRA
DATA ANALYSTS
MANUFACTURING DOWDTIME

Machine Learning Model



Benefits:

- Predicts delays before they happen
- Enhances production planning and scheduling accuracy
- Identifies high-risk operators and batches
- Reduces unplanned downtime and improves line efficiency
- Supports data-driven decision making at the factory floor

Planning Accuracy

- Daily downtime forecasting improves scheduling.
- Estimated reduction in unplanned delays: ~12–15%.
- Most productive day: [most_productive_day], helps optimize batch assignments.

Early Risk Detection

- Flags high-risk operators and lines.
- Top operators by batch count: [top 3 operators].
- Early alerts prevent large cycle time deviations.

Maintenance Prioritization

- Focus on most impactful downtime factors.
- Top factors by incidents: [list top factors].
- Enables proactive maintenance, reducing reactive fixes.

Analysis Automation

- Saves hours of manual reporting.
- Interactive dashboards visualize: operator downtime, product-size effects, temporal patterns, and 12-factor breakdown.

Business Impact

Challenges

Data Quality:

- Inconsistent formats and occasional errors in recording batch start/end times, affecting analysis and prediction accuracy.

Imbalanced Data:

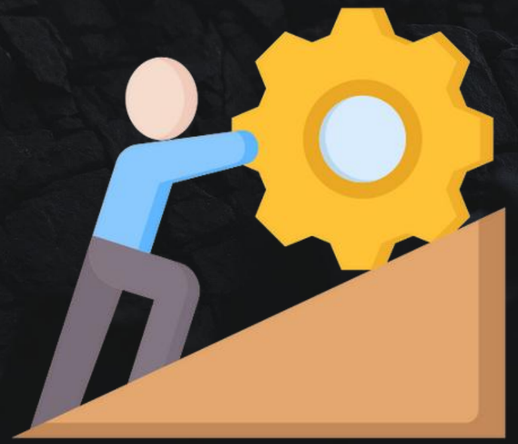
- Majority of batches have zero downtime, while a few have significant downtime.
- Makes regression model training challenging, as large downtime values are rare compared to zeros.

Lack of Real-Time Data:

- All data is post-event logs rather than live sensor streams.
- Limits the ability to detect high-risk batches or lines in real time.

Limited Historical Data (Data Set) :

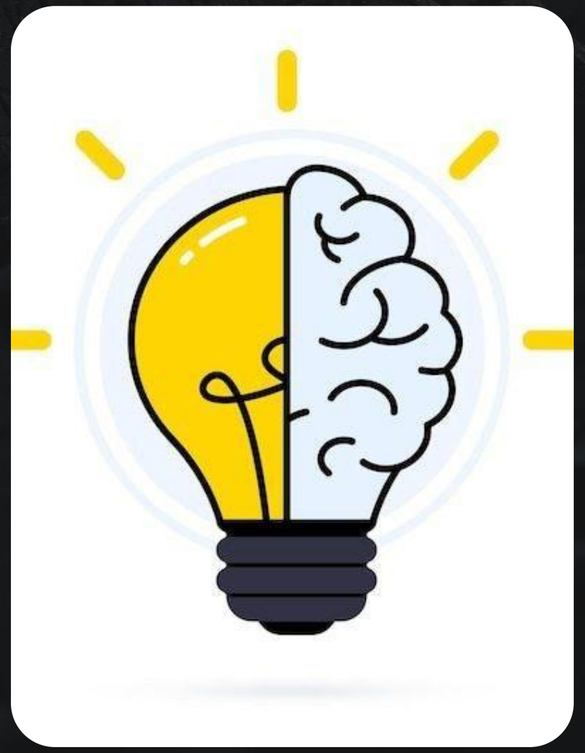
- Available historical records cover a restricted time period.
- Reduces the model's ability to learn rare downtime patterns.
- More historical data would improve predictive accuracy.



QUANTRA
DATA ANALYSTS
MANUFACTURING DOWDTIME

To Solve these problems:

Solutions



1 **Robust Cleaning Pipeline:**

Implementing specific logic to handle overnight times and standardize text-based data (Operators, Sizes).

2 **Advanced Feature Engineering:**

Creation of aggregated statistical features (e.g., historical average downtime per operator) to boost prediction accuracy.

3 **Strong Model Selection:**

Utilizing robust tree-based algorithms (like XGBoost) for their resilience and ability to handle feature complexity.

4 **Integrated Deployment:**

Transitioning the project from a notebook to a practical tool via API and an interactive dashboard.

QUANTRA
DATA ANALYSTS
MANUFACTURING DOWDTIME

Limitations

Disadvantages

The project covers a limited historical period, which may not capture rare downtime events, causing model accuracy to drop on unseen or unusual failure modes. Data quality issues, such as inconsistent start/end times, and reliance on post-event logs instead of real-time sensors, can affect predictions. Manual operator inputs introduce potential human error and variability, while the high prevalence of zero-downtime batches makes predicting extreme cases more challenging.

Limitations

Disadvantages

Real-World Operational Challenges:

In practice, production lines face unexpected machine breakdowns, operator fatigue, and supply chain disruptions that are not captured in historical data. Coordination between maintenance, operations, and quality teams can also cause delays and reduce overall efficiency.

Recommendations

Efficient and Sustainable Recommendations

- Implement preventive maintenance schedules guided by model insights to reduce unplanned downtime.
- Provide operator training programs focused on high-risk tasks identified by the predictive model.
- Establish standardized data collection processes for more consistent and high-quality inputs.
- Continuously monitor model performance and update the system based on real-world deviations.

Conclusion

Finally

This project showcases the true power of combining data, machine learning, and operational knowledge. By meticulously analyzing historical production and downtime data, we built a predictive model that estimates total downtime per batch, helping operators and managers act proactively.

Beyond the model, we developed an interactive dashboard to visualize trends, monitor key metrics, and deliver actionable insights. To further enhance operational support, we also implemented a chatbot capable of answering queries and guiding decisions in real-time.

This project reflects not only the power of our predictive models and dashboards but also the dedication and teamwork behind every insight. It showcases how our combined effort transforms raw production data into actionable, real-world solutions.

QUANTRA
DATA ANALYSTS
MANUFACTURING DOWDTIME

Thank You