

# Data and AI Made Useful for the Front-Line

The information operators need to be high performers, available in real-time.

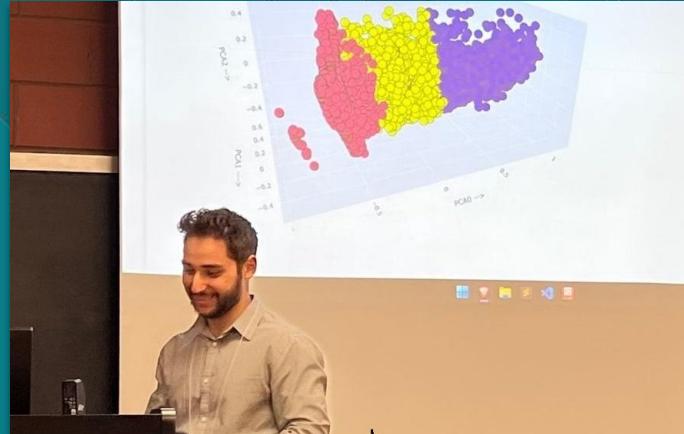


# Joe Youssouf

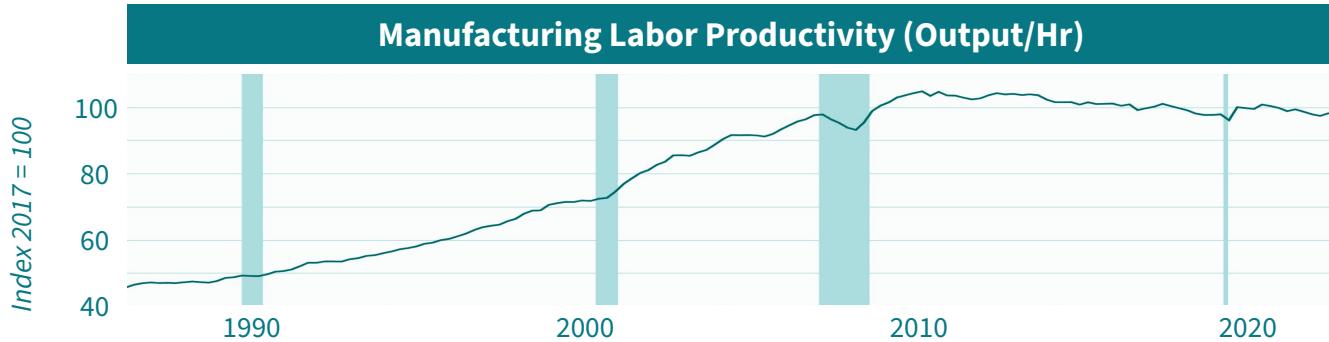
Data Scientist II  
Oden Technologies

[joseppy.ca](http://joseppy.ca)

<https://github.com/JYoussouf/slides/gdg-devfest-2024>



# A Perfect Storm for Manufacturers



A Typical Manufacturer's Workforce		
	<u>2018</u>	<u>2024</u>
Average Tenure	20 Years	3 Years
Average Time in Position	7 Years	9 Months
Average 3-Month Retention Rate	90%	50%

- ✖ Productivity Plateauing
- ✖ Workforce Challenges
- ✖ Demand & Asset Growth
- ✖ Margin & Competitive Pressures



# Front-Line Challenges



Dependent on Operator Experience



Decreasing Tenure



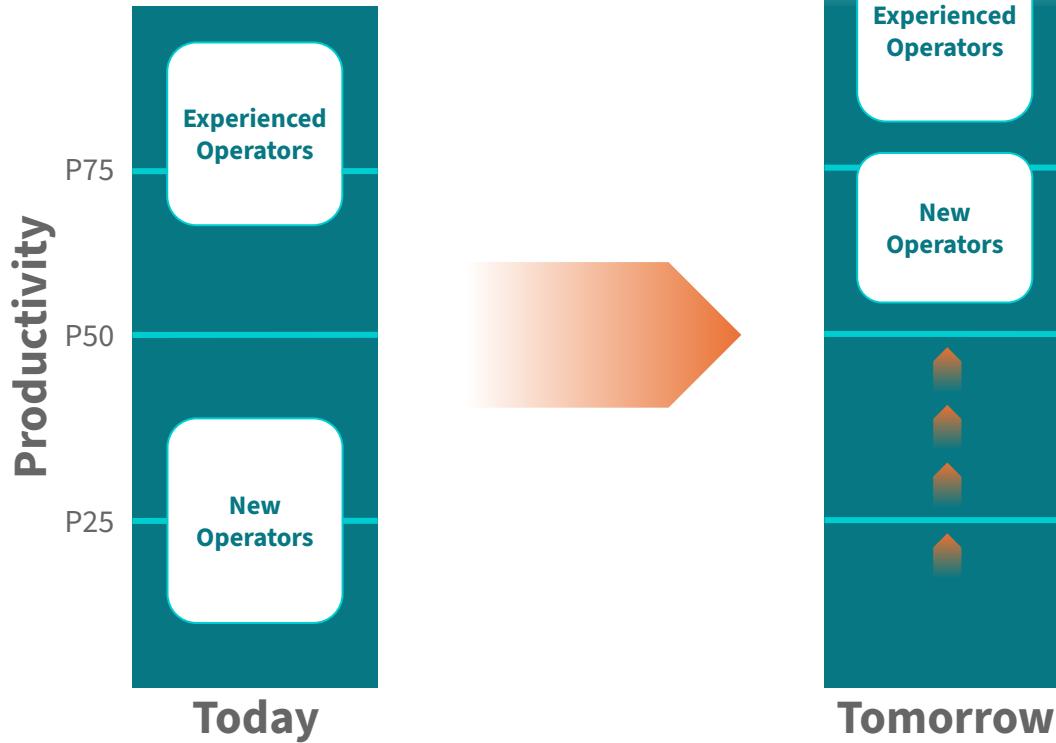
Losing Tribal Knowledge





# Empower the Front-Line Real-Time AI Recommendations

# Bridging the Productivity Gap

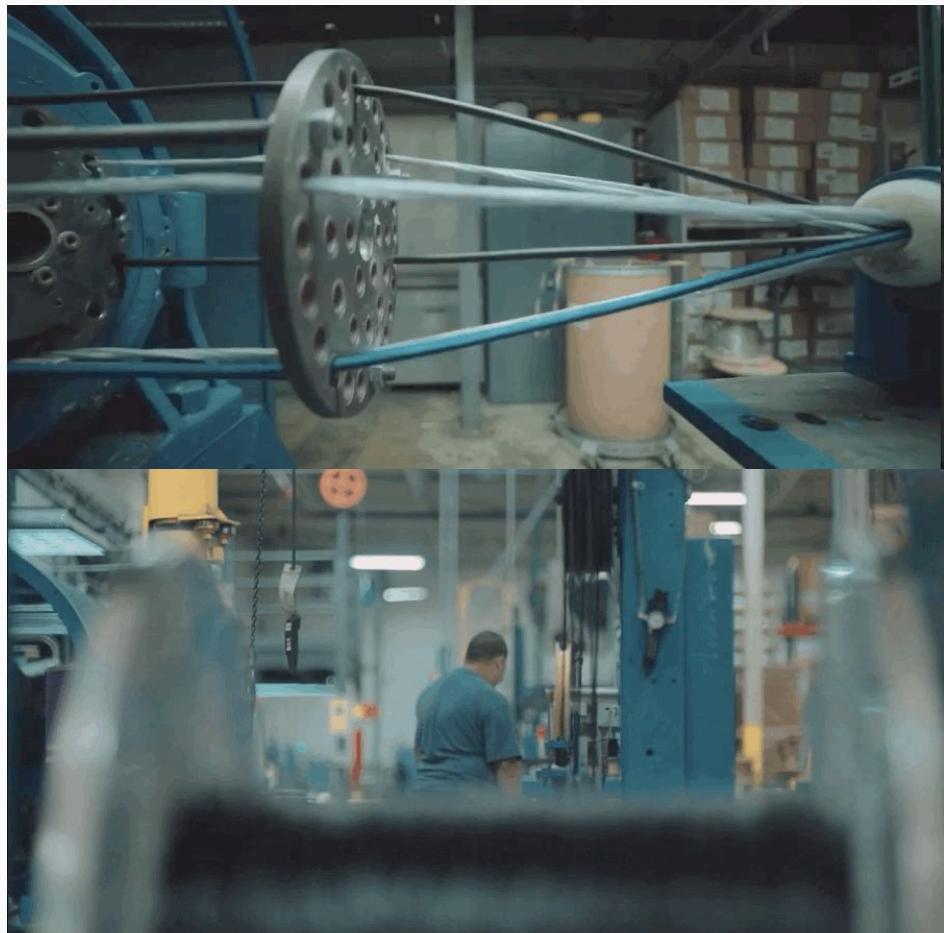


# Oden's Customers

Medium to large manufacturers in plastics extrusion, injection molding, pipes, chemical, paper and pulp.

Process and Quality Engineers looking to centralize, analyze, and act on their data.

Plant managers who are looking to optimize logistics, output, and cost.



# The Oden Product Suite



## Data Engine

Data Cleansing,  
Contextualization,  
Inferencing & Enrichment



## Process AI

AI-Enabled Predictions  
AI-Enabled Recommendations



## Knowledge AI

AI-Enabled Work Instructions  
Operator CoPilot

*Pre-release*



## Factory Analytics

Ad-Hoc Reporting & Dashboarding  
Data Analytics & Visualization



## Data Connect

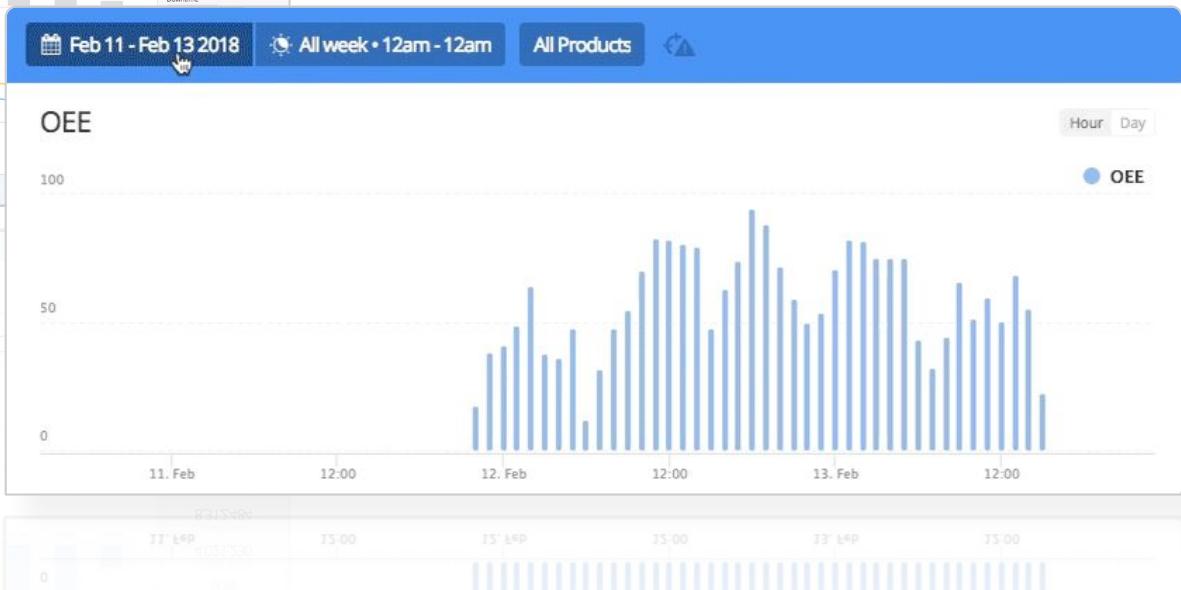
Live & Batch Data Export

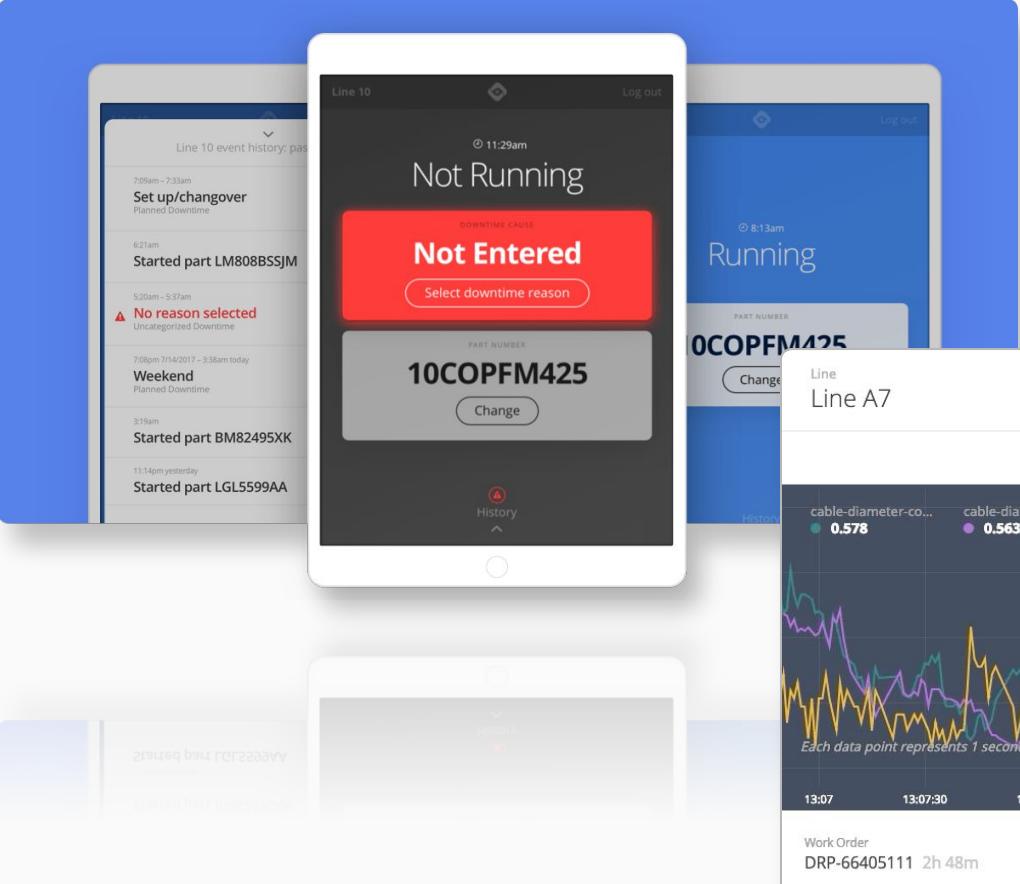
*Pre-release*

# Interactive Time-series Analysis



- Compare performance across different equipment.
- Visualize hourly uptime and key custom metrics.
- Calculations for analyzing and optimizing factory performance.





## Real Time Manufacturing Data

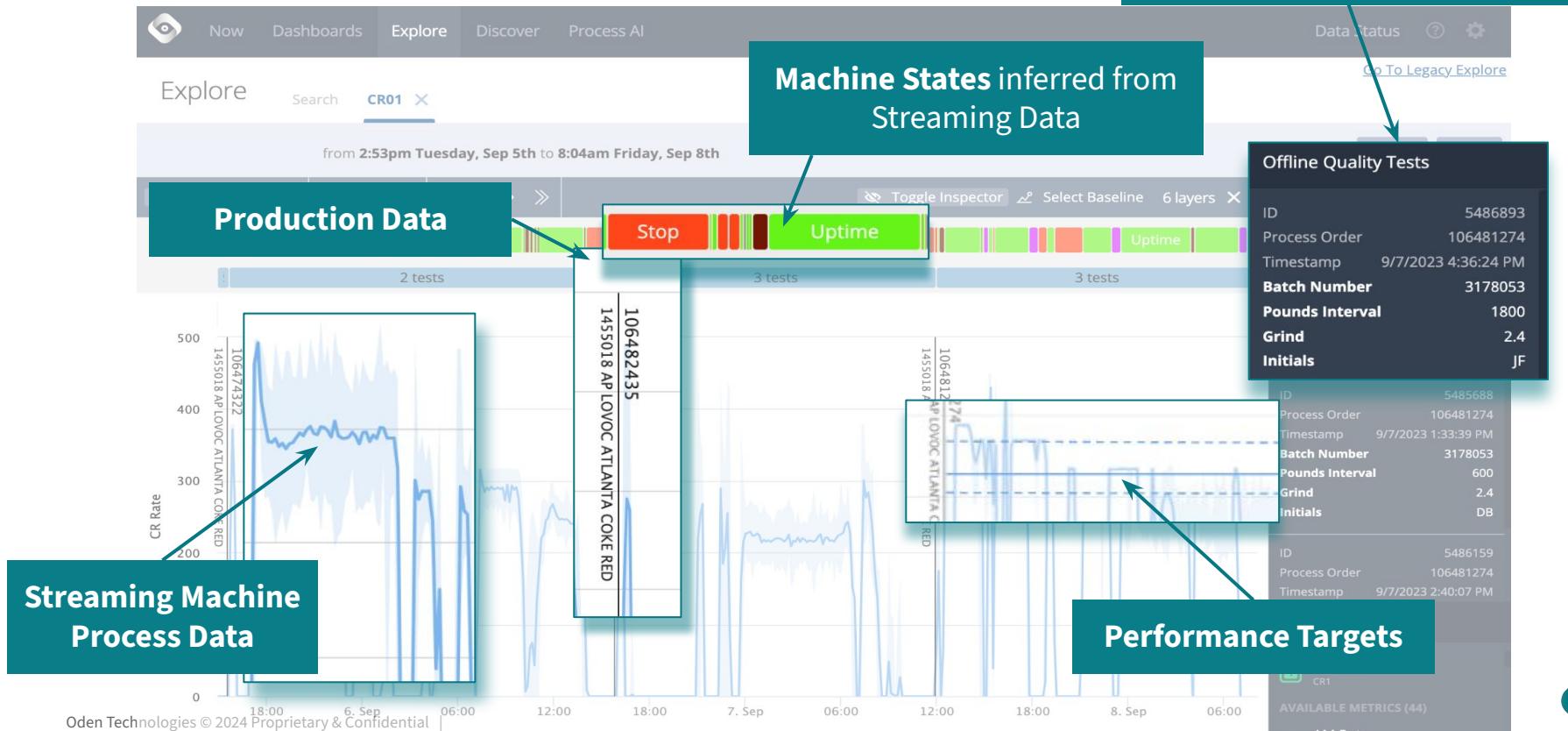
- Streaming second-by-second metrics
- Interactive app that prompts on production state changes and collects user input.



# Data Contextualization

# Oden Data Engine

# Offline Quality Tests



# Process AI: Real-Time Operator Recommendations

Now Dashboards Explore Discover Labs Process AI Data Status ⓘ ⚙️

Process AI

Engineer  Operator

LINE Paper Machine 02

PRODUCT 6528C962

Stable periods (10/12/2023 - 4/12/2024)  
Select a point below to view the process settings for the given speed and cost

COST

\$0.25/ft  
\$0.20/ft  
\$0.15/ft  
\$0.10/ft

LINE SPEED (FPM)

1000 1100 1200

● Stable period  
● Current period  
● Recommended period

LINE SPEED (FPM)

1,088.20 > 1,106.34

MATERIAL FLOW 2D

CURRENT 780.52 > RECOMMENDED 834.70

REFINER FLOW 2D

CURRENT 1,109.53 > RECOMMENDED 921.47

REFINER FLOW 2E

CURRENT 731.31 > RECOMMENDED 688.54

STEAM FLOW 1A

CURRENT 84.91 > RECOMMENDED 80.28

PAPER STRENGTH TENSILE

OFFLINE 354.05 PREDICTED 347.10 TARGET 330.00

MOISTURE CONTENT

OFFLINE 12.75 PREDICTED 12.62 TARGET 12.50

PAPER STRENGTH Z PLY

OFFLINE 178.35 PREDICTED 183.87 TARGET 180.00

Real-time predicted quality metrics

Overall goals and progress

CURRENT COST \$0.1772/ft

PREDICTED COST \$0.1418/ft

PREDICTED SAVINGS \$2,344.59/hr

Start Run

Oden Technologies © 2024 Proprietary & Confidential

# Process AI: Incremental Recommendations for Paper

Now Dashboards Explore Discover Labs Process AI Data Status ⓘ ⚙️

Process AI

Configure

GRADE CHANGE → MOISTURE → Z PLY → COST

LINE: Paper Machine 02 | PRODUCT: 6528C962

Stable periods (3/15/2023 - 8/1/2023)  
Select a point below to view its process settings

77.4000  
77.3000  
77.2000  
77.1000  
77.0000  
76.9000

COST (\$/FT)

375 400 425

LINE SPEED (FPM)

CURRENT: 380.21 > RECOMMENDED: 399.93

MATERIAL FLOW 2D

CURRENT: 316.96 ✓ RECOMMENDED: 316.30

REFINER FLOW 2D

CURRENT: 345.52 » RECOMMENDED: 402.94

PAPER STRENGTH Z PLY

OFFLINE: 80.228 PREDICTED: 82.71 TARGET: 132

84  
83  
82

08:30 09:00 09:30 10:00

Predicted

PAPER STRENGTH TENSILE

OFFLINE: 147.543 PREDICTED: 144.65 TARGET: 132

146  
145  
144

08:30 09:00 09:30 10:00

Predicted

MOISTURE CONTENT

OFFLINE: 5.102 PREDICTED: 5.051 TARGET: 5

5.0512  
5.05115  
5.0511

08:30 09:00 09:30 10:00

Predicted

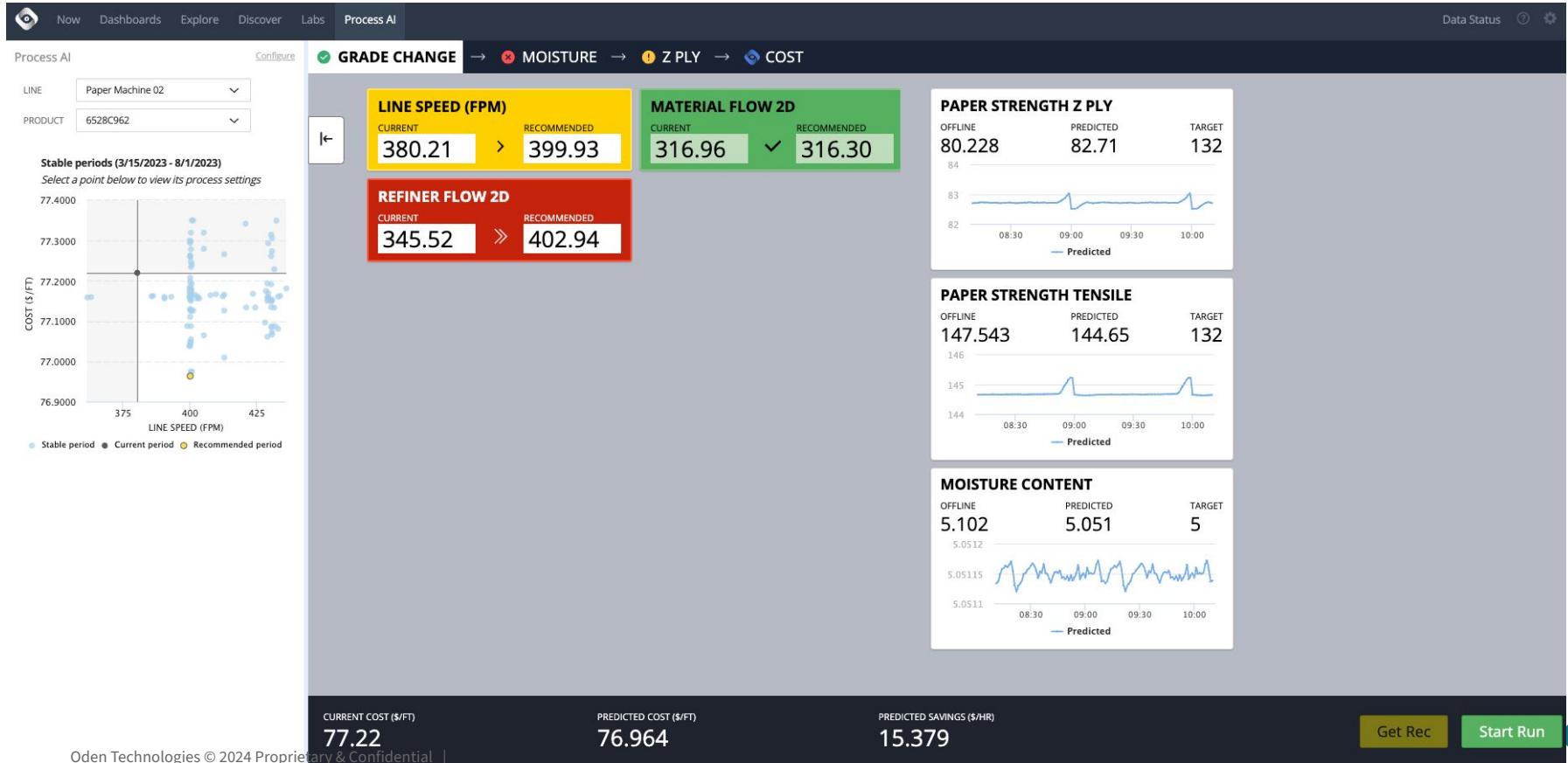
CURRENT COST (\$/FT): 77.22

PREDICTED COST (\$/FT): 76.964

PREDICTED SAVINGS (\$/HR): 15.379

Get Rec Start Run

Oden Technologies © 2024 Proprietary & Confidential |



## **Process A1: Incremental Recommendations for Paper**

The screenshot displays a process monitoring interface for a paper machine. At the top, a navigation bar includes 'Now', 'Dashboards', 'Explore', 'Discover', 'Labs', and 'Process AI'. On the far right, there are 'Data Status' and user profile icons. The main area is titled 'Process AI' and shows a flowchart: GRADE CHANGE → MOISTURE → Z PLY → COST.

**Left Panel:**

- Line:** Paper Machine 02
- Product:** 6528C962
- Stable periods:** Select a point below to view its process settings. A graph shows MOISTURE levels (5.0000 to 5.1250) over LINE SPEED (FPM) (400 to 450). It indicates 'Determining current conditions...' and highlights three periods: Stable period (light blue), Current period (dark blue), and Recommended period (yellow).

**Middle Panel:**

- MATERIAL FLOW 2D:** CURRENT 316.86, RECOMMENDED 345.52
- LINE SPEED (FPM):** CURRENT 380.14, RECOMMENDED 465.87
- REFINER FLOW 2D:** CURRENT 346.29, RECOMMENDED 358.42
- REFINER FLOW 2E:** CURRENT 287.98, RECOMMENDED 343.29
- STEAM FLOW 1A:** CURRENT 34.75, RECOMMENDED 34.95

**Right Panel:**

- PAPER STRENGTH Z PLY:** OFFLINE 80.228, PREDICTED 82.71, TARGET 132. A line chart shows predicted values from 08:30 to 10:00.
- PAPER STRENGTH TENSILE:** OFFLINE 147.543, PREDICTED 144.65, TARGET 132. A line chart shows predicted values from 08:30 to 10:00.
- MOISTURE CONTENT:** OFFLINE 5.102, PREDICTED 5.051, TARGET 5. A line chart shows predicted values from 08:30 to 10:00.

**Bottom Panel:**

- CURRENT COST (\$/FT):** 77.22
- PREDICTED COST (\$/FT):** -
- PREDICTED SAVINGS (\$/HR):** -

**Buttons:** Get Rec, Start Run

Oden Technologies © 2024 Proprietary & Confidential |

# Process AI: Incremental Recommendations for Paper

Now Dashboards Explore Discover Labs Process AI Data Status ⓘ ⚙

Process AI

Configure

LINE Paper Machine 02

PRODUCT 6528C962

Stable periods (3/15/2023 - 8/1/2023)  
Select a point below to view its process settings

77.4000  
77.3000  
77.2000  
77.1000  
77.0000  
76.9000

77.4000  
77.3000  
77.2000  
77.1000  
77.0000  
76.9000

375 400 425

LINE SPEED (FPM)

● Stable period ● Current period ○ Recommended period

GRADE CHANGE → MOISTURE → Z PLY → COST

MATERIAL FLOW 2D

CURRENT 316.79 ✓ RECOMMENDED 316.30

LINE SPEED (FPM)

CURRENT 380.27 > RECOMMENDED 399.93

REFINER FLOW 2D

CURRENT 345.14 >> RECOMMENDED 402.94

REFINER FLOW 2E

CURRENT 287.70 > RECOMMENDED 307.52

STEAM FLOW 1A

CURRENT 34.83 ✓ RECOMMENDED 31.55

PAPER STRENGTH Z PLY

OFFLINE 80.228 PREDICTED 82.71 TARGET 132

84  
83  
82

08:30 09:00 09:30 10:00

Predicted

PAPER STRENGTH TENSILE

OFFLINE 147.543 PREDICTED 144.65 TARGET 132

146  
145  
144

08:30 09:00 09:30 10:00

Predicted

MOISTURE CONTENT

OFFLINE 5.102 PREDICTED 5.051 TARGET 5

5.0512  
5.05115  
5.0511

08:30 09:00 09:30 10:00

Predicted

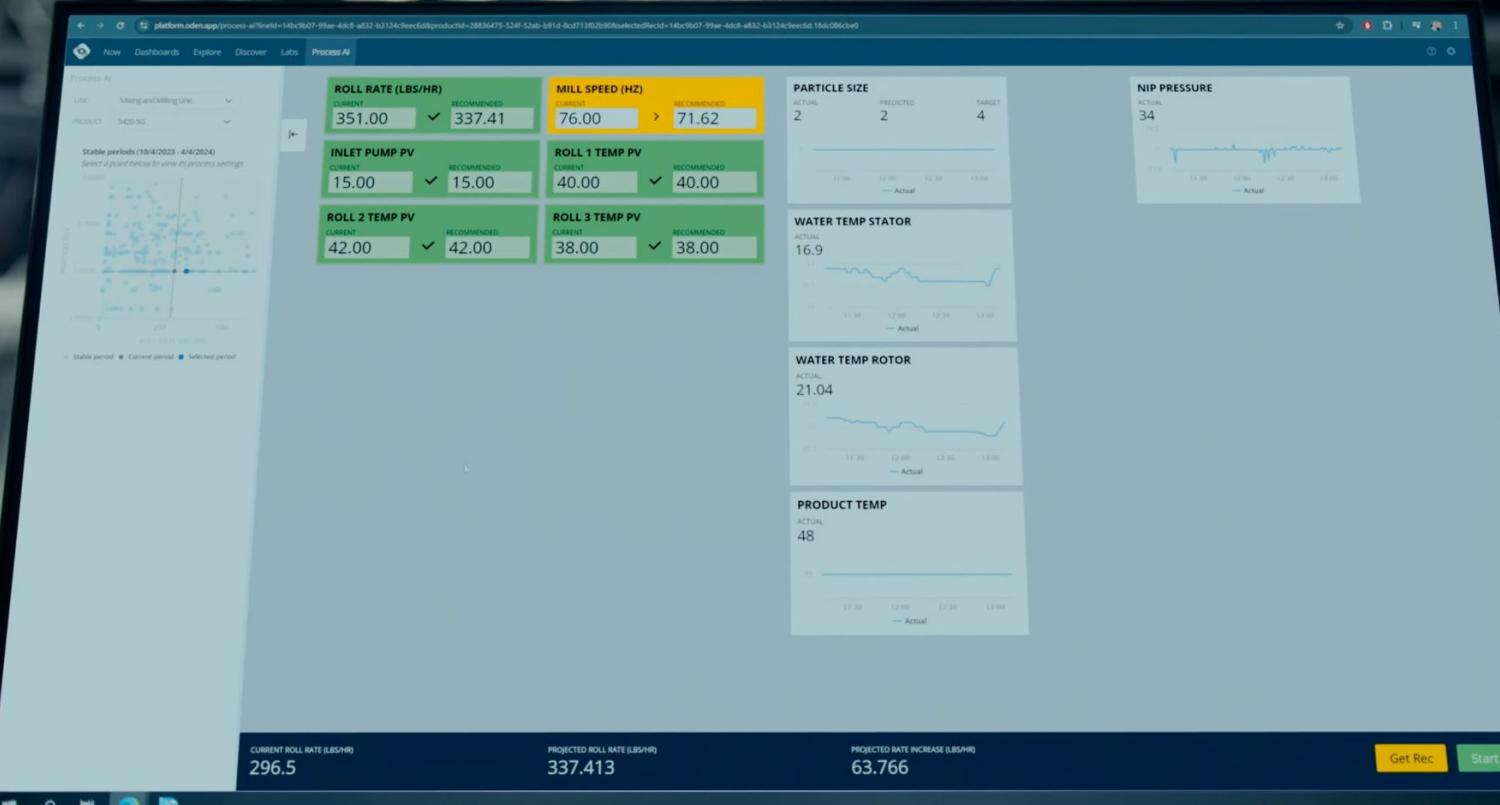
CURRENT COST (\$/FT) 77.22

PREDICTED COST (\$/FT) 76.964

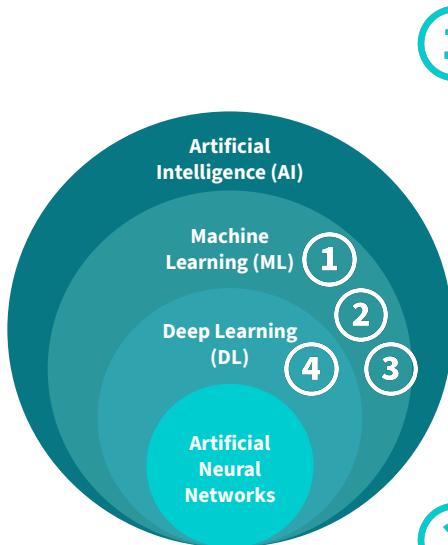
PREDICTED SAVINGS (\$/HR) 15.379

Get Rec Start Run

Oden Technologies © 2024 Proprietary & Confidential |



# AI/ML Techniques used in the Oden Product Suite



## 1 Linear Models

Uses: Predictive Quality

Benefits: Interpretability, fast, easy to train, scalable

Oden uses a variety of different approaches (LASSO, Elastic Net, PCA-Regression) to predict quality metrics. These are powerful as all of the model variables can be visualized and trust is built.

Oden also uses Deep Learning and other models for PQ problems to validate results.

## 2 Tree Models and Ensembles

Uses: Predictive Quality, Root Cause Analysis

Benefits: Interpretability, non-linear separations, handling categorical data

Used to identify key contributors separating periods of interest from baselines, i.e. changeover detection

## 3 Optimization Models

Uses: Process AI Recommendations

Benefits: Solve complex multi-objective multi-constraint problems

Oden uses optimization models to provide process setting recommendations directly to operators in real-time.

## 4 Deep Learning and Large Language Models

Uses: Stable Periods, Knowledge AI [in development]

Benefits: Powerful learning

Oden uses Deep Learning models for Stable Period analysis for Process AI, as well as in exploratory work with Knowledge AI and providing work instructions.



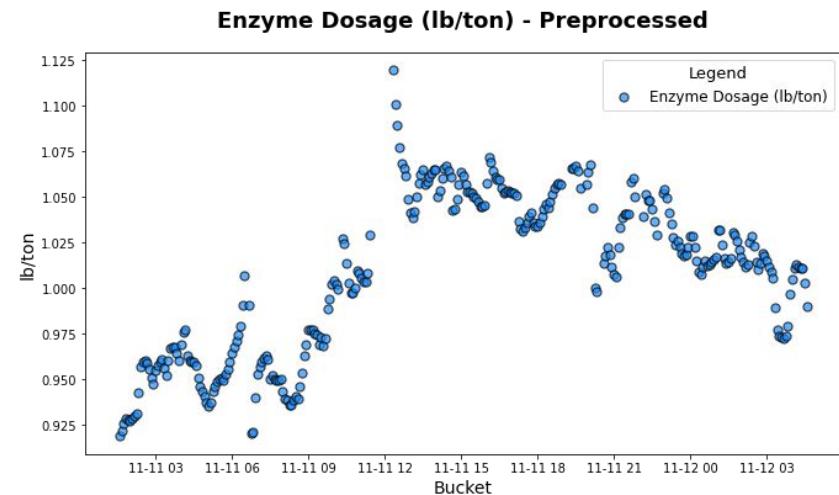
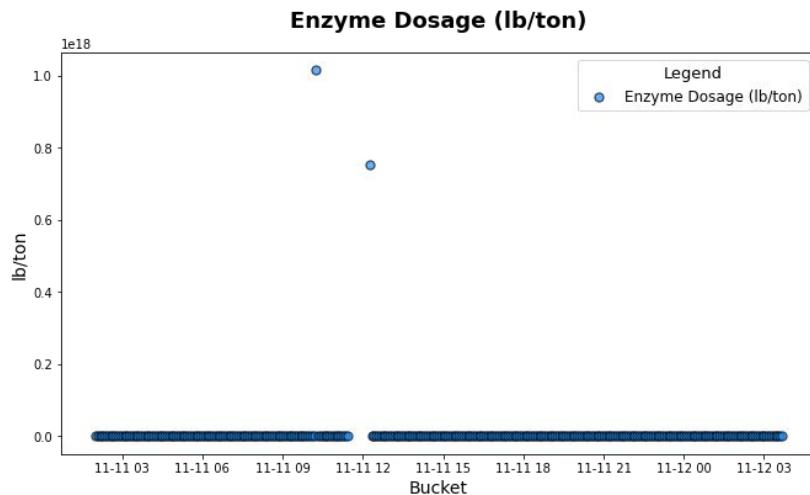
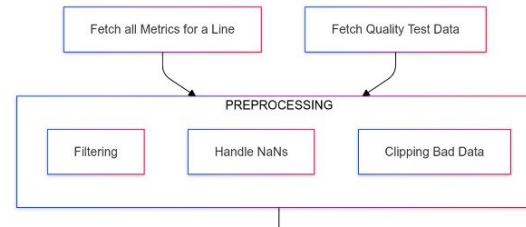
# Problem Statement: Predictive Quality

Use multi-dimensional time series signatures:

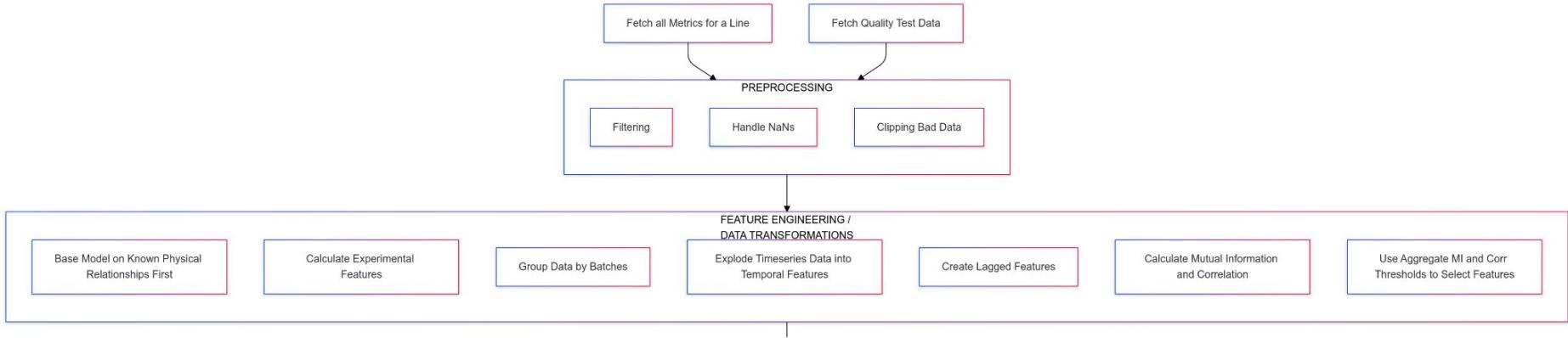
To predict a single quality test **label** at the end of each run:



# Example PQ Pipeline - Preprocessing

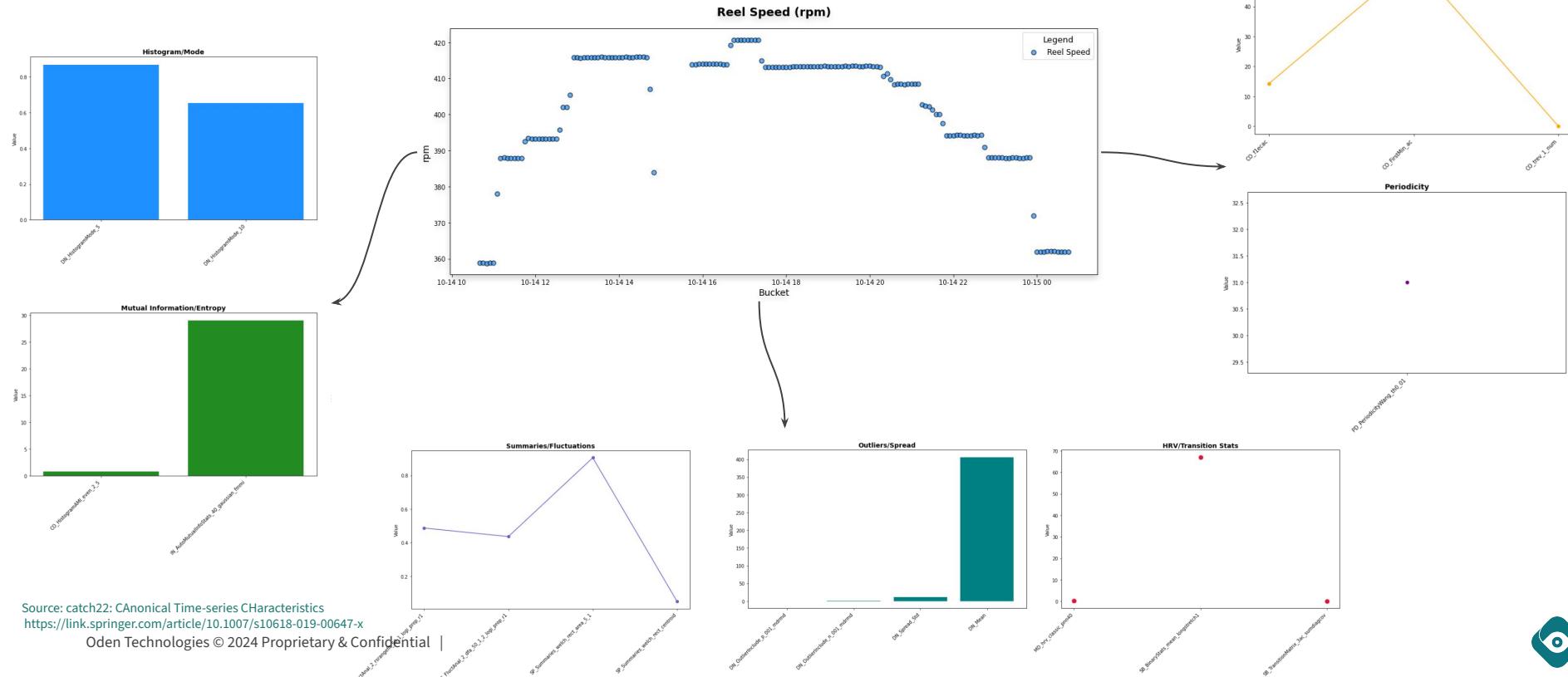


# Example PQ Pipeline - Feature Engineering



# Example PQ Pipeline - Feature Engineering

## Time Series Decomposition into Core Stats



Source: catch22: Canonical Time-series Characteristics  
<https://link.springer.com/article/10.1007/s10618-019-00647-x>

Oden Technologies © 2024 Proprietary & Confidential |



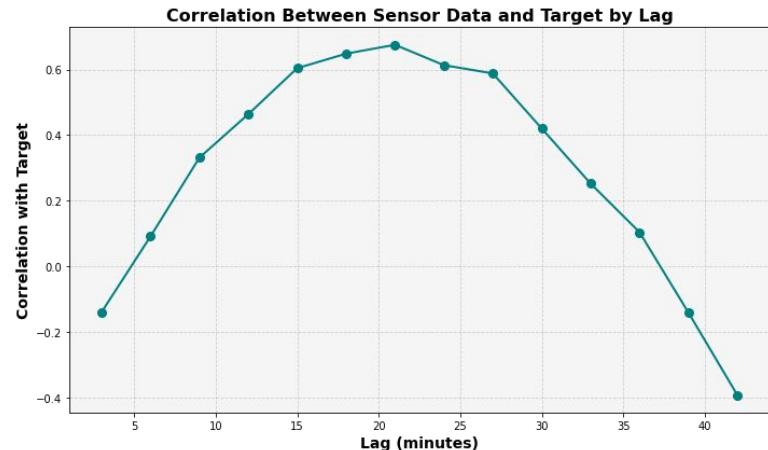
# Feature Selection and the Lag Effect

A change in one metric may not impact output quality for N minutes.

How can we find N?

If available, ask a Domain Expert!

Otherwise, we can estimate this with lagged features, and ranking metrics (e.g. corr)



bucket	cd821ae1-c69b-5d37-b751-84c28bcf14e9	cd821ae1-c69b-5d37-b751-84c28bcf14e9_lag_5min	cd821ae1-c69b-5d37-b751-84c28bcf14e9_lag_20min	cd821ae1-c69b-5d37-b751-84c28bcf14e9_lag_60min
2024-04-09 19:30:00+00:00	2.053912	1.997217	2.514325	2.594364
2024-04-09 19:35:00+00:00	2.027539	2.065962	2.516855	2.779457
2024-04-09 19:40:00+00:00	2.014745	2.091880	2.469408	2.942021
2024-04-09 19:45:00+00:00	2.015056	2.080691	2.381821	2.937846
2024-04-09 19:50:00+00:00	1.985830	2.084497	2.358898	2.894803
...	...	...	...	...
2024-11-14 23:35:00+00:00	3.318202	3.758257	3.936180	3.781135
2024-11-14 23:40:00+00:00	3.298853	3.529197	3.918642	3.784988
2024-11-14 23:45:00+00:00	3.275909	3.468211	3.874753	3.764397
2024-11-14 23:50:00+00:00	3.304411	3.396445	3.864001	3.851068
2024-11-14 23:55:00+00:00	3.321725	3.347016	3.862047	3.946134
31806 rows × 4 columns				



# Feature Selection and Ranking Metrics

## Feature Space:

-> 100's to 1000's of raw & transformed metrics

## Systematic Filtering Methods:

1. Human-In-The-Loop Feedback (Physics-Based/ State-Space Modelling)

2. Mutual Information/ Absolute Correlation Scores

3. Concentration of Noise/“Bad Data”, nunique(), etc.

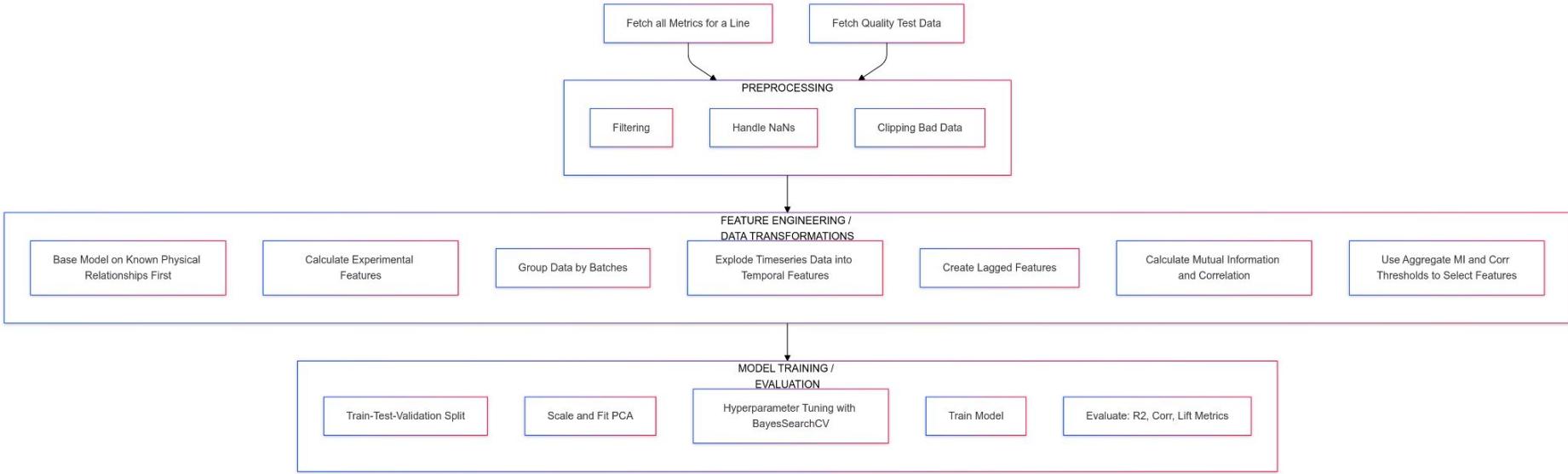
metric	mutual_information	abs_correlation	unique_values
10119216-9108-48b8-835b-29a83ae4bc5f-CO_Histog...	0.007816	0.020257	352
671b338c-b982-4dbe-adb0-393126f8f64c-CO_FirstM...	0.028868	0.082565	11
b6c70614-6e25-4a7b-86f7-a02702fedcf7-FC_LocalS...	0.028162	0.037197	10
d6786dea-bf97-43d0-83d3-60d9e1b0c88e-DN_Histog...	0.000168	0.004669	1378
experiment_7-FC_LocalSimple_mean3_stderr	0.000000	0.002482	1378
...	...	...	...
306f4d01-8f33-5c89-ad05-11bf51b3b050-MD_hrv_cl...	0.000000	0.022463	20
1d2e853d-f2e5-58e2-af68-baaf57d4fa12-CO_EMBED2...	0.012775	0.006904	1378
e2644829-1edb-50e9-8f20-4a3c97788227-DN_Spread...	0.000000	0.034116	1378
b3a5ca35-34e2-5eaf-83df-85617f630f0e-SB_Binary...	0.000603	0.080383	10
b6c70614-6e25-4a7b-86f7-a02702fedcf7-CO_Histog...	0.000563	0.005759	327

$$I(X, Y) = - \sum_{x_i, y_i} p(x_i, y_i) \log_2 \frac{p(x_i, y_i)}{p(x_i)p(y_i)}.$$

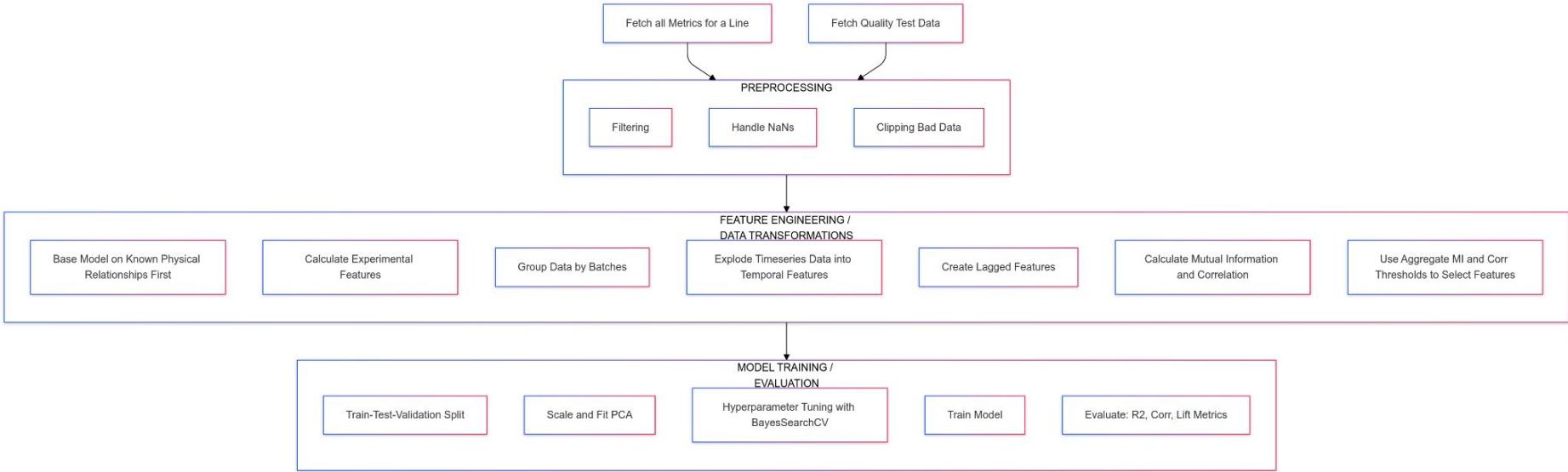
```
-> from sklearn.feature_selection import  
    mutual_info_regression  
-> pd.Series.corr(labels)  
-> pd.Series.nunique()
```



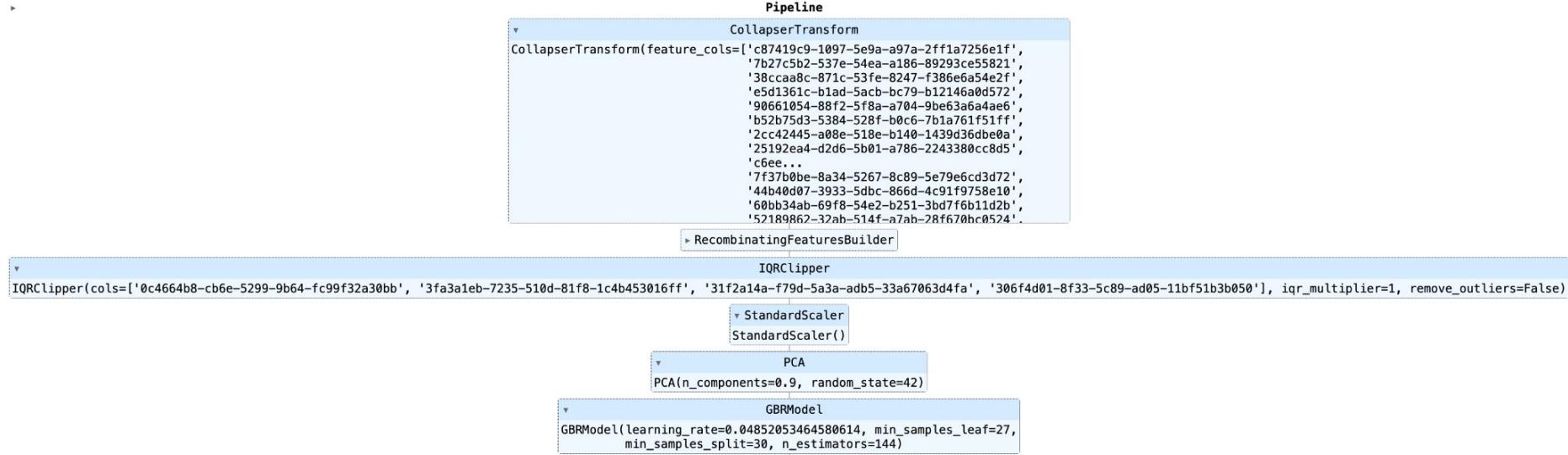
# Example PQ Pipeline - Train/Deploy



# Example PQ Pipeline - Train/Deploy



# Example Pipeline - Registered to MLFlow



# MLFlow Artifacts

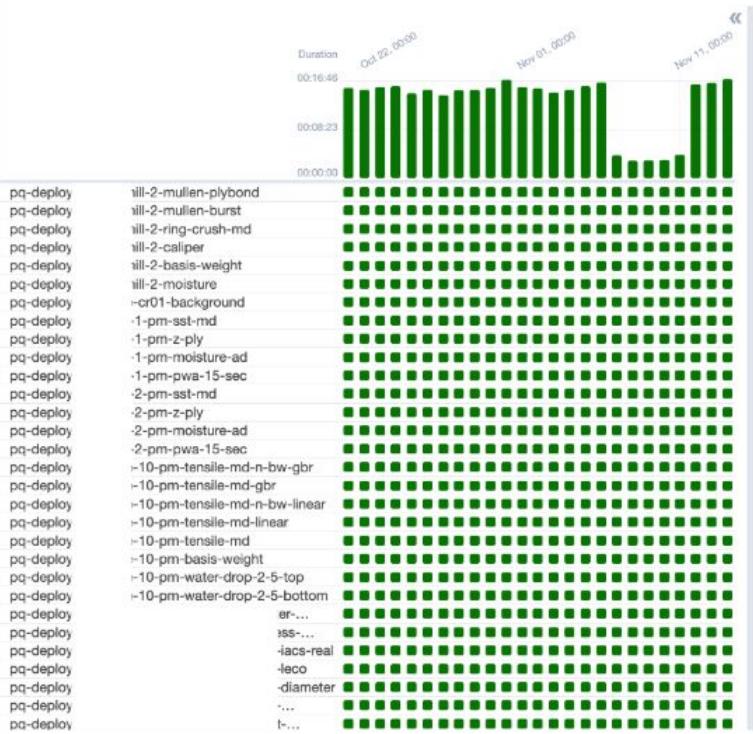
PQ\_v2\_All\_Features\_PCA\_Tuned

Overview   Model metrics   System metrics   **Artifacts**

- ▼ **data\_split**
  - ☒ test\_split.csv
  - ☒ train\_split.csv
- ▼ **model**
  - **metadata**
    - 📄 MLmodel
    - 📄 conda.yaml
    - 📄 model.pkl
    - 📄 python\_env.yaml
    - 📄 requirements.txt
  - ▼ **model\_outputs**
    - ☒ model\_predictions.csv
  - **model\_params**
    - 📄 dataflow\_config.json
    - 📄 inference\_metadata.json
    - 📄 pipeline\_vis.html
    - 📄 pq\_training.log



# Model Deployment via Apache Airflow



## DAG predictive\_quality\_deploy\_v1

**Details**   **Graph**   **Gantt**   **Code**

### DAG Runs Summary

Total Runs Displayed	25
<span style="color: green;">■ Total success</span>	25
First Run Start	2024-10-22, 00:00:01 UTC
Last Run Start	2024-11-15, 00:00:01 UTC
Max Run Duration	00:16:46
Mean Run Duration	00:12:58
Min Run Duration	00:02:54

### DAG Summary

Total Tasks	37
MLFactoryCloudRunJobOperators	35
OdenGKEPodOperator	1
PassiveCheckOperator	1

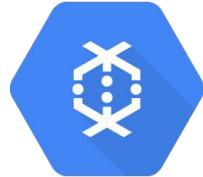
### DAG Details

Dag id	predictive_quality_deploy_v1
Description	null



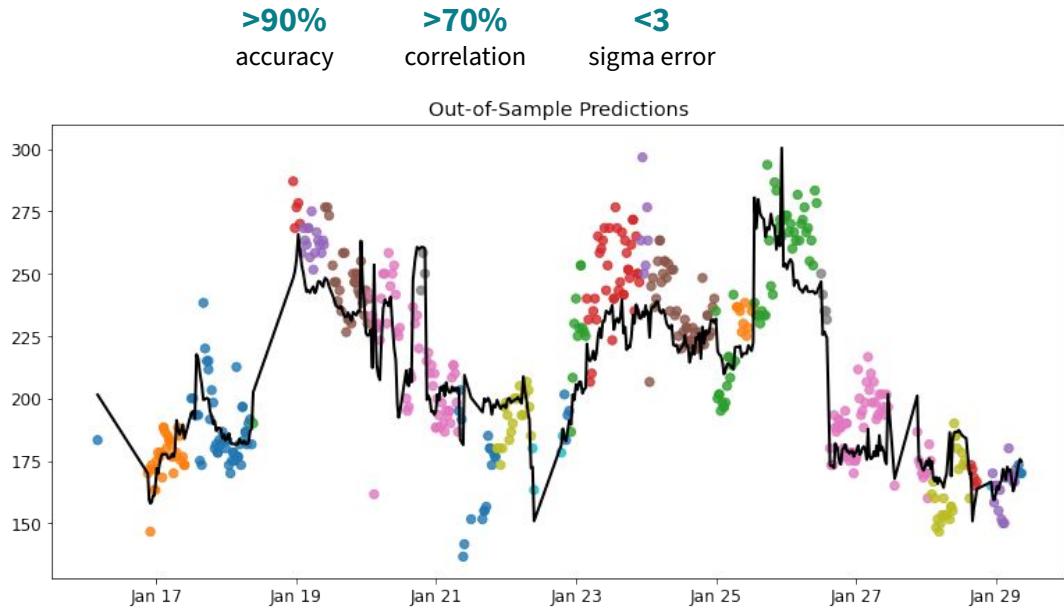
# Real-Time Predictions with Dataflow (Google Cloud)

- DataFlow job manages real-time inference/ posting predictions to the Oden Platform



# Key Variables Example: Paper

## Predictive Quality Model Development



Ply Bond - Paper	
Key Variable	Coefficient
#7 Vat Flow	-6.5224
#10 Vat Flow	-4.1779
Shared Waste Water	-4.0135
2nd Press Bottom Felt Vacuum	-3.5552
Bone Dry Weight (scanner)	-3.3621
Production Rate (tons/hr)	-2.8738
D Refiner HPDT	2.2074
1st Dryer Section Speed	2.988
C Refiner KW	3.1313
C Refiner HPDT	3.1394
% Target	3.5729
#1 Vat Lvl Output	3.5746
D Refiner KW	4.0107
Total HPDT	8.8416

Total Process Variables: 127

Model Variables: 38

Key Variables: 14



# Tangible Results for Pulp & Paper

Fortune 500 Paper Customer

**Goal:** Optimize speed & reduce input Costs while maintaining quality

**Solution:** Oden Process AI + Data Engine, first 6 months

**3%** Increase to Line Speed  
(\$440K benefit/year)

**5%** Reduction in Material &  
Energy Costs  
(\$505K benefit/year)



**\$900K+**  
Savings / Year



# Tangible Results for Wire & Cable

**Customer:** #1 Electrical cable producer in North America

**Goal:** Increase production through higher line speeds

**Solution:** Oden Process AI

**When:** 90 days from go-live

**Scope:** 2 extrusion lines, many products

**Who:** 20 Operators of varying experience

**Results:**

+20%

Increase in  
Production Rate

+5.3M<sub>ft</sub>

Additional Feet  
Produced over 90  
days





# Thank You!

[oden.io](https://oden.io)

[joseppy.ca](https://joseppy.ca)

