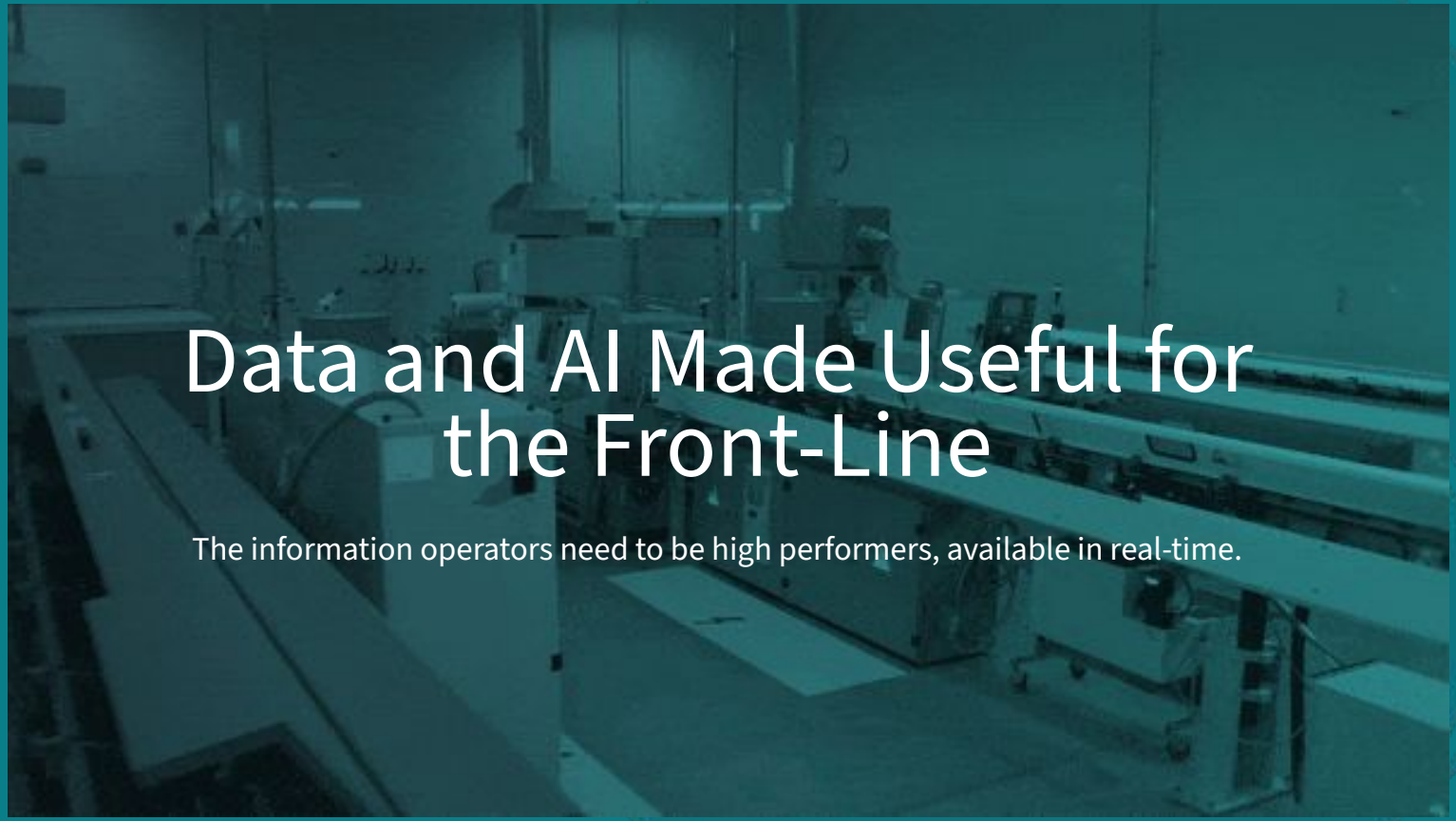


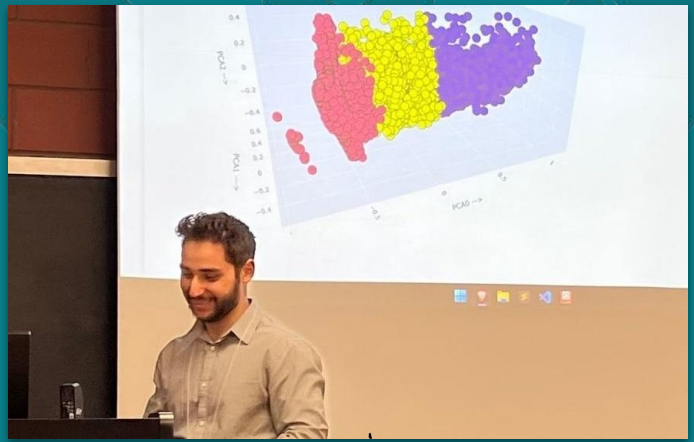
Data and AI Made Useful for the Front-Line

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

Joe Youssouf

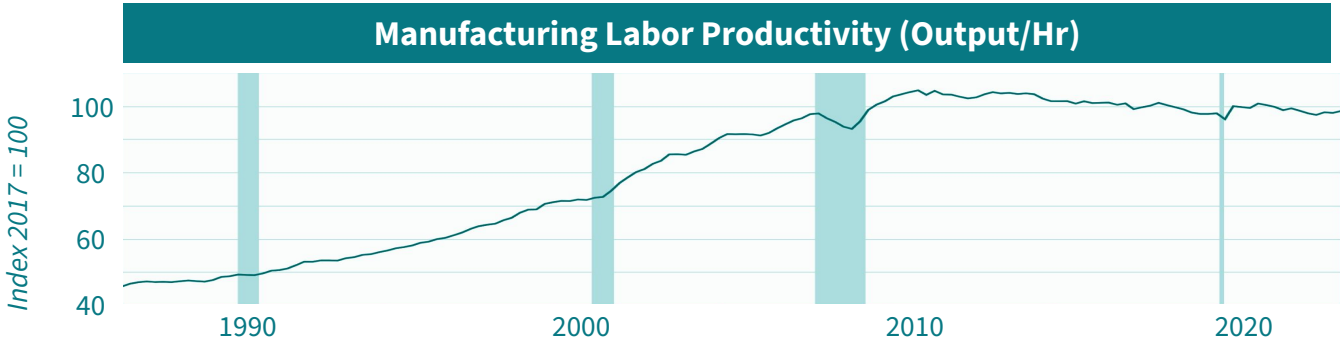
Data Scientist II
Oden Technologies

joseppy.ca



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

A Perfect Storm for Manufacturers



A Typical Manufacturer's Workforce		
	2018	2024
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

Sources: LNS Research, FRED, U.S. Bureau of Labor Statistics

Front-Line Challenges

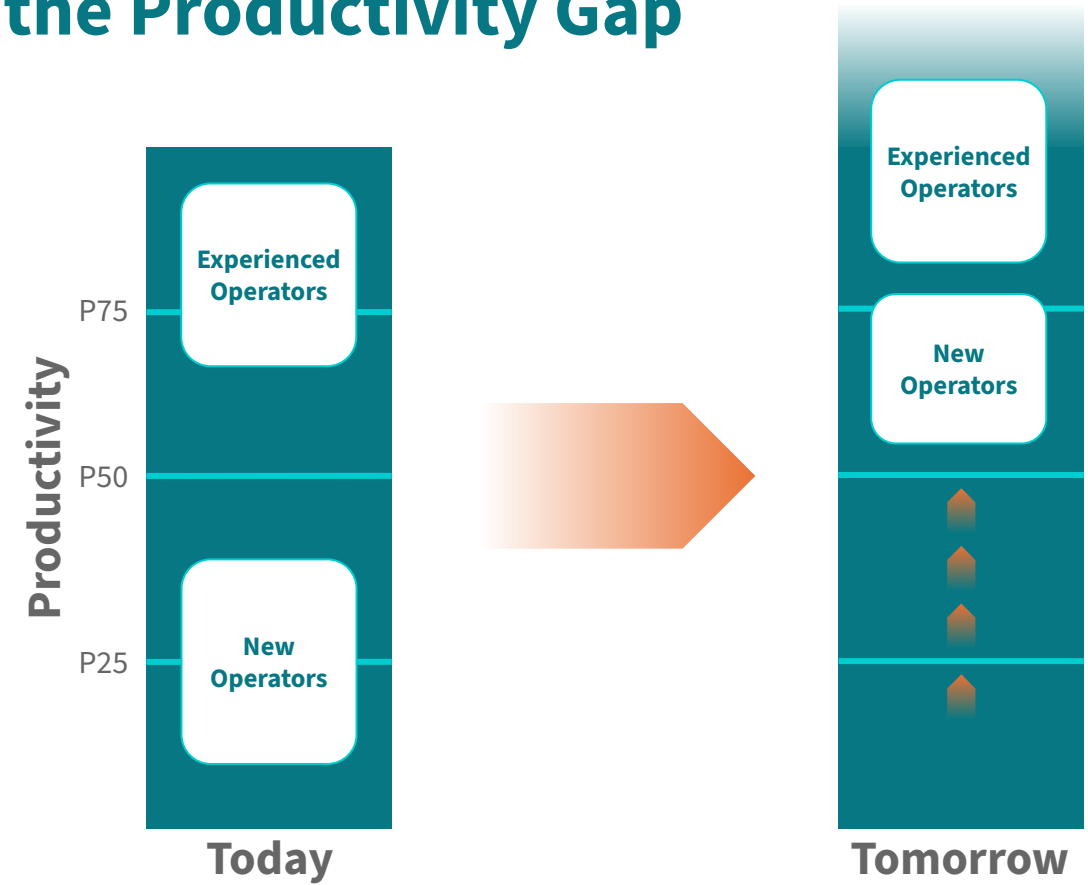
- Dependent on Operator Experience
- Decreasing Tenure
- Losing Tribal Knowledge

 Sources: LNS Research

ODEN TECHNOLOGIES

Empower the Front-Line Real-Time AI Recommendations

Bridging the Productivity Gap

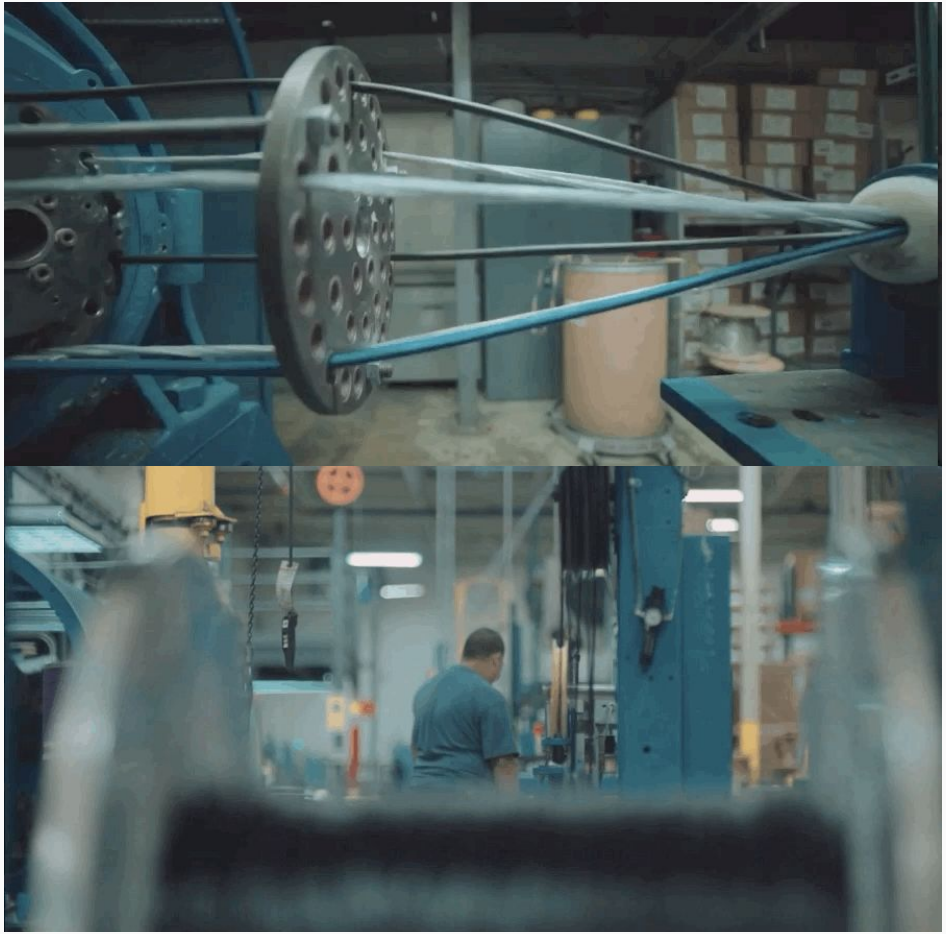


Oden's Customers

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
Pre-release
AI-Enabled Work Instructions
Operator CoPilot

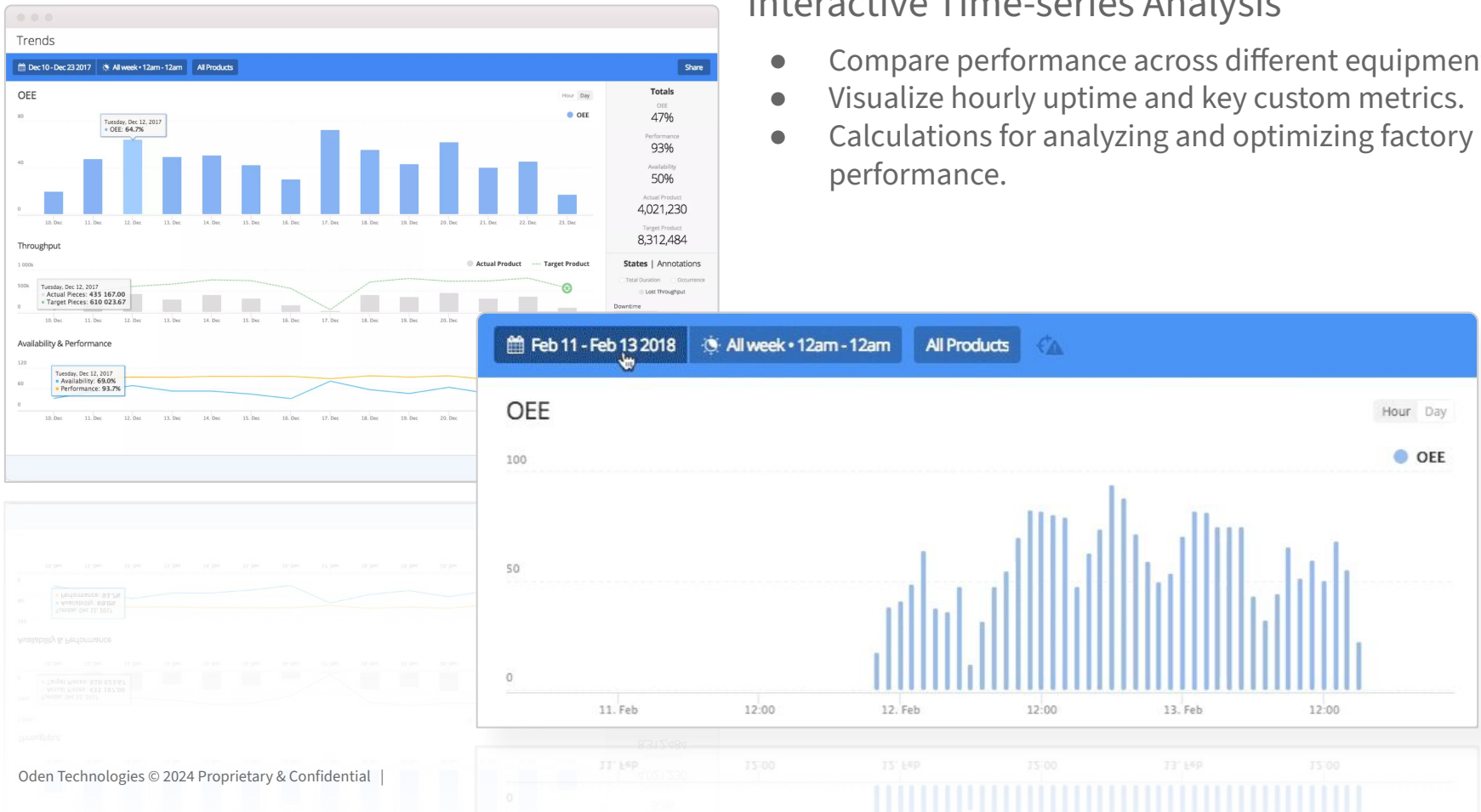
Factory Analytics
Ad-Hoc Reporting & Dashboarding
Data Analytics & Visualization

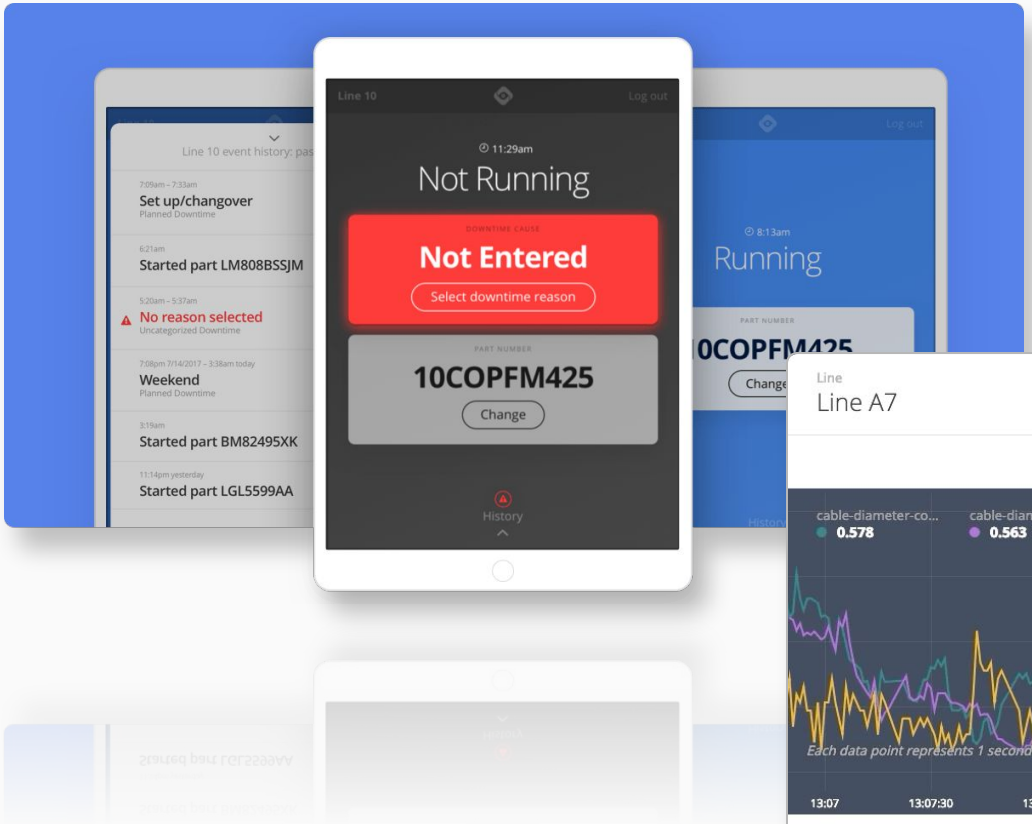
Data Connect
Pre-release
Live & Batch Data Export

ODEN TECHNOLOGIES

Interactive Time-series Analysis

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



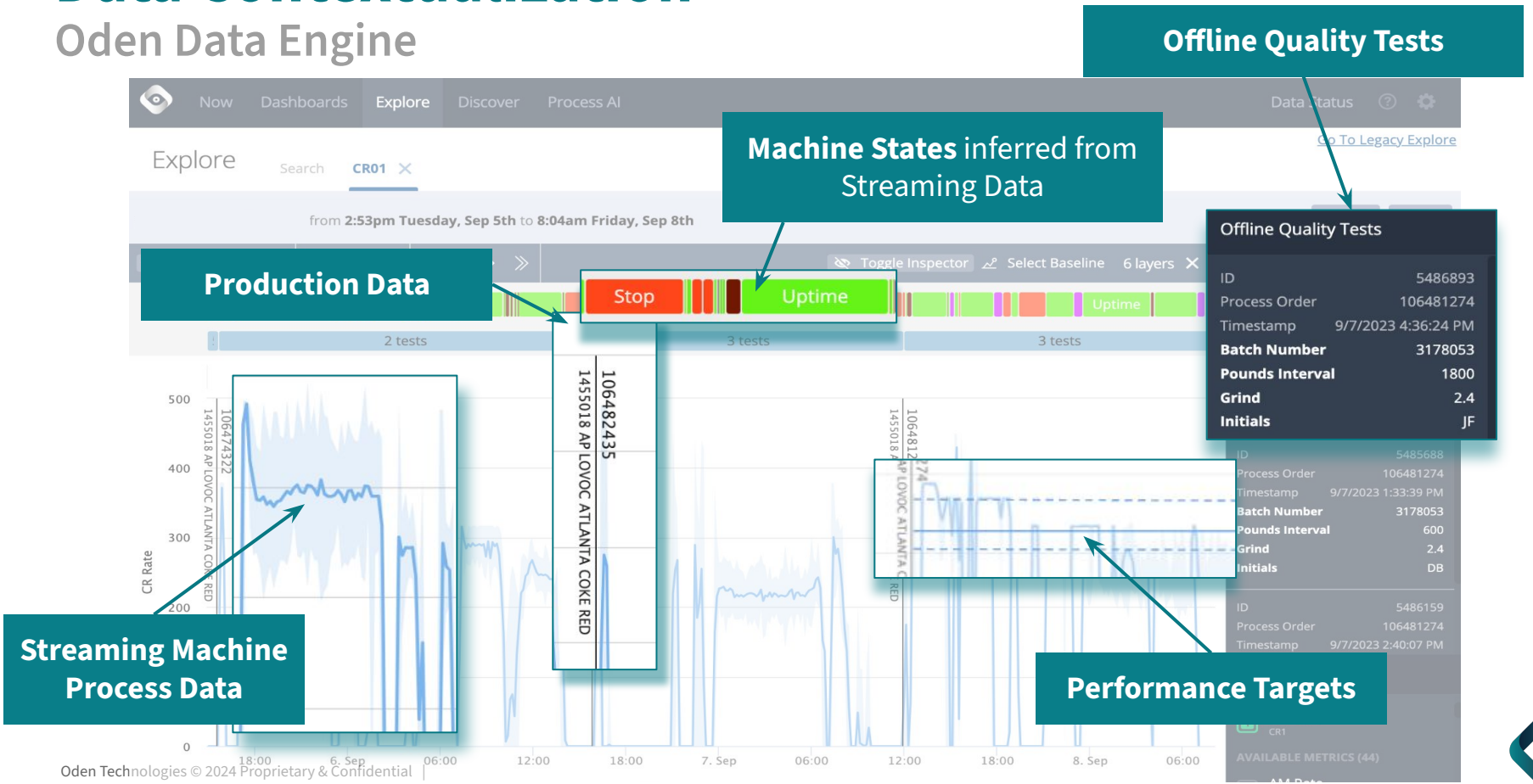


Real Time Manufacturing Data

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

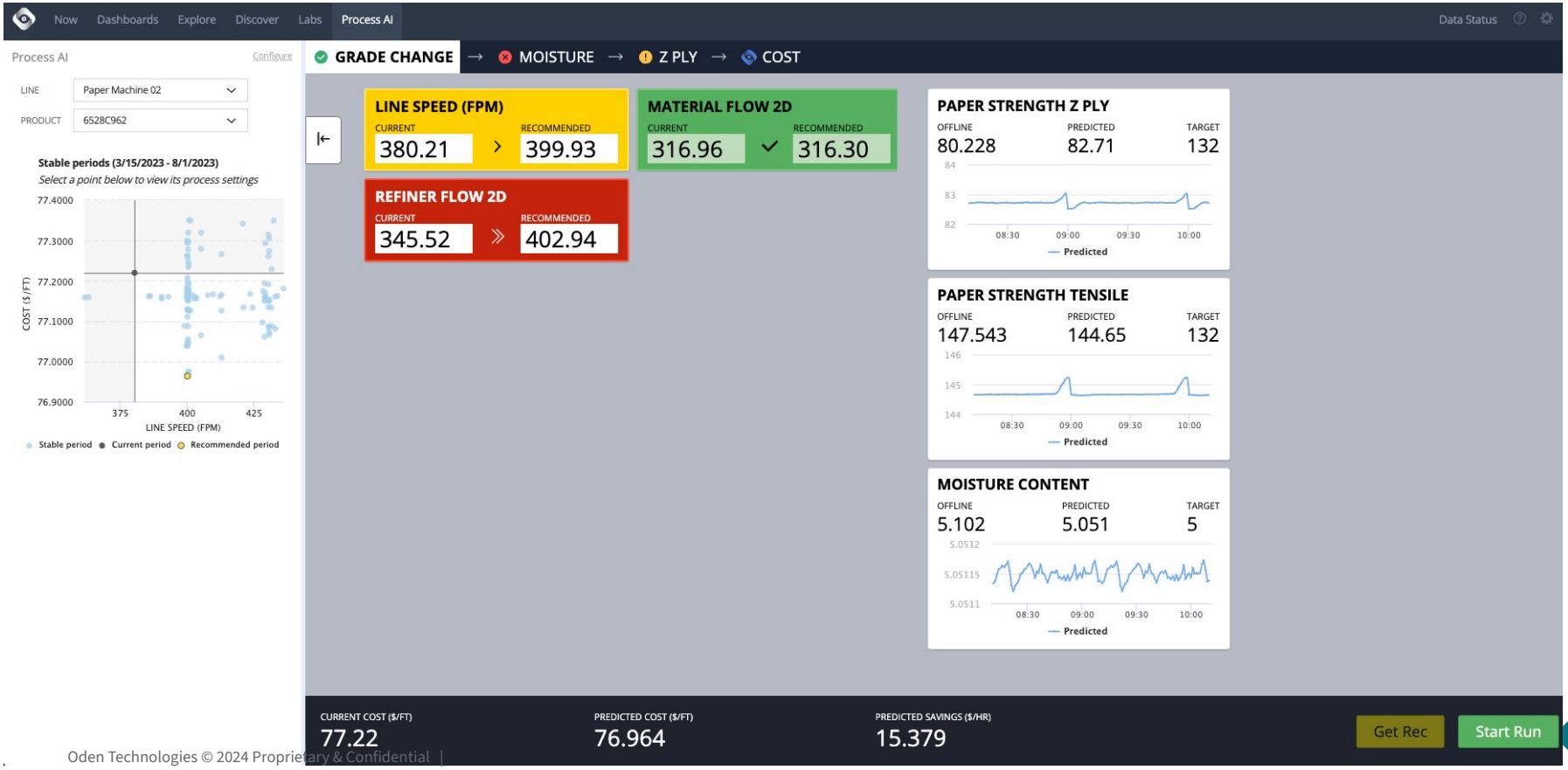
Data Contextualization

Oden Data Engine

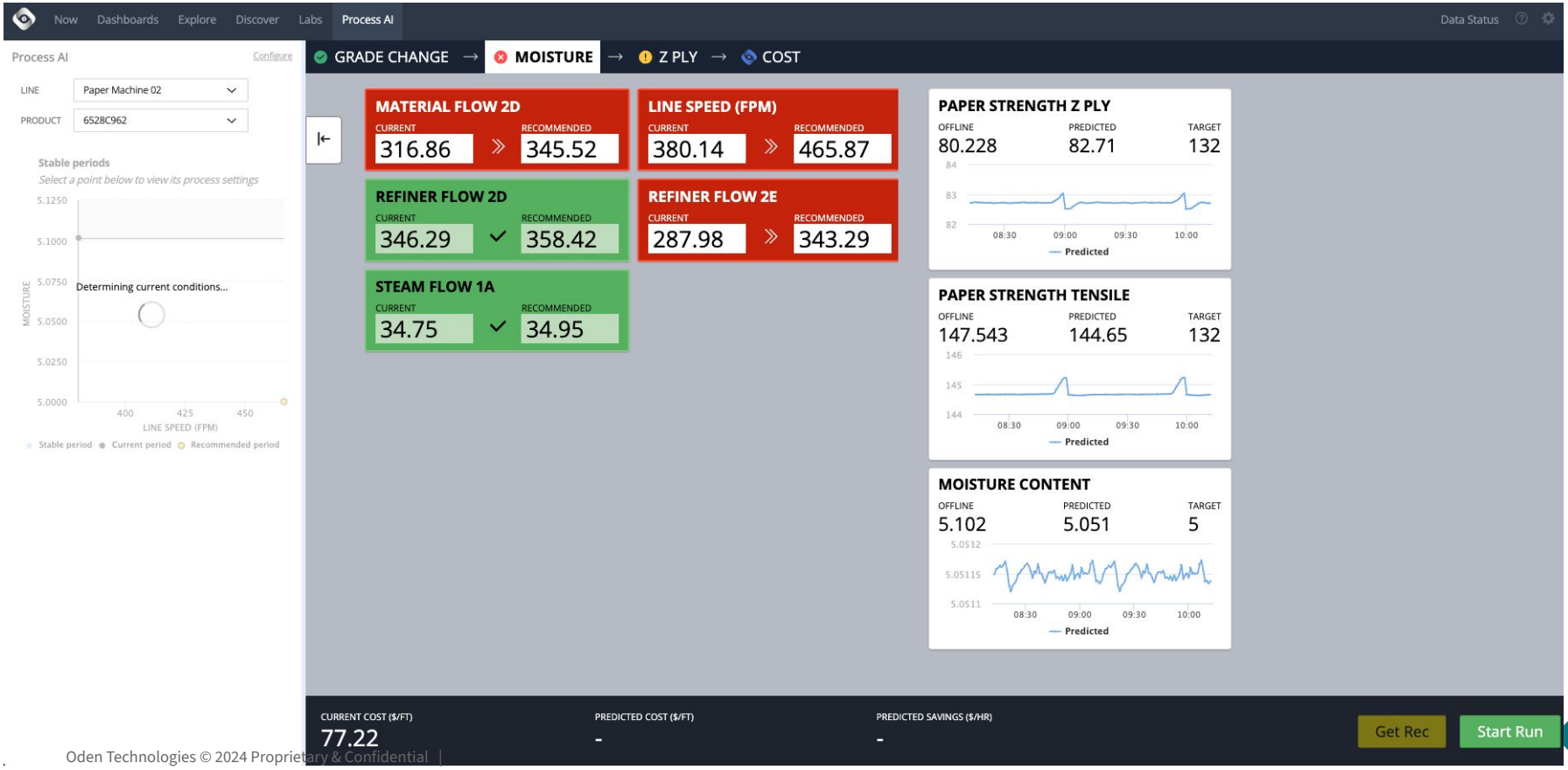


Process AI: Real-Time Operator Recommendations

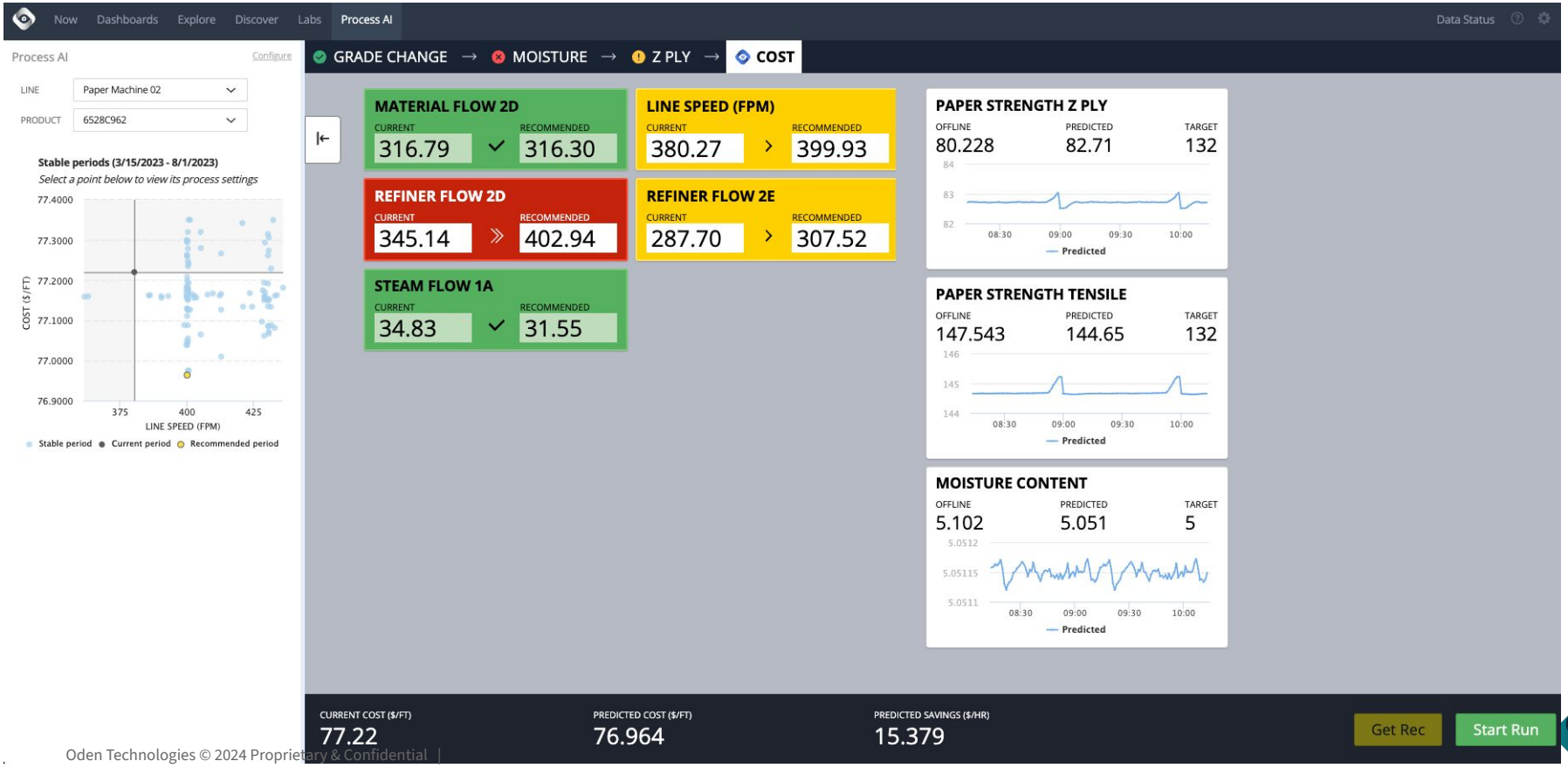
Process AI: Incremental Recommendations for Paper

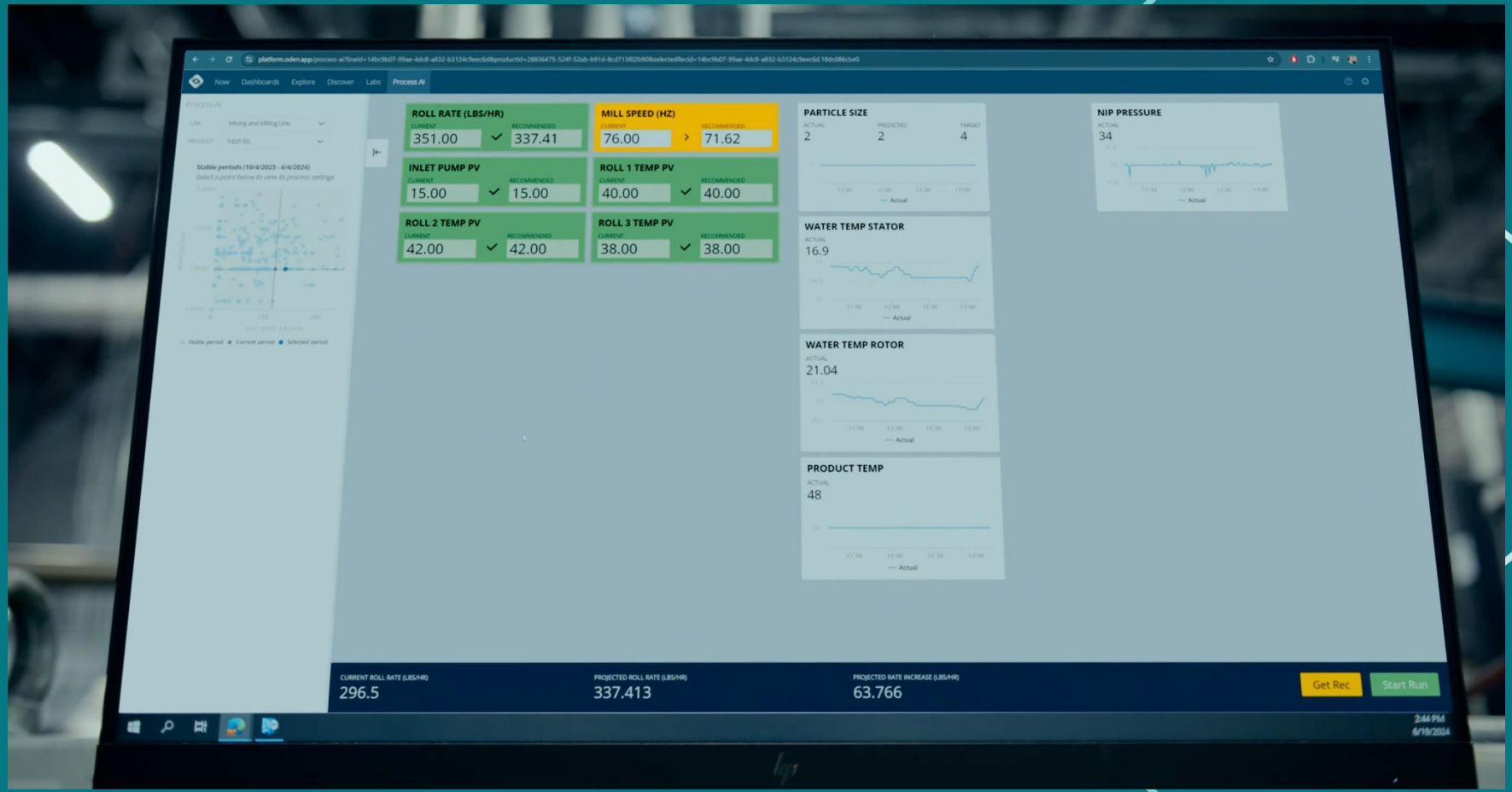


Process AI: Incremental Recommendations for Paper

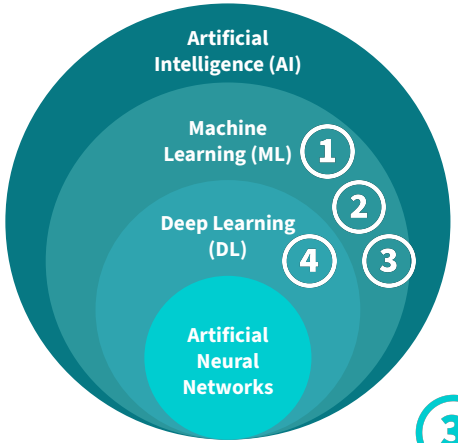


Process AI: Incremental Recommendations for Paper





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

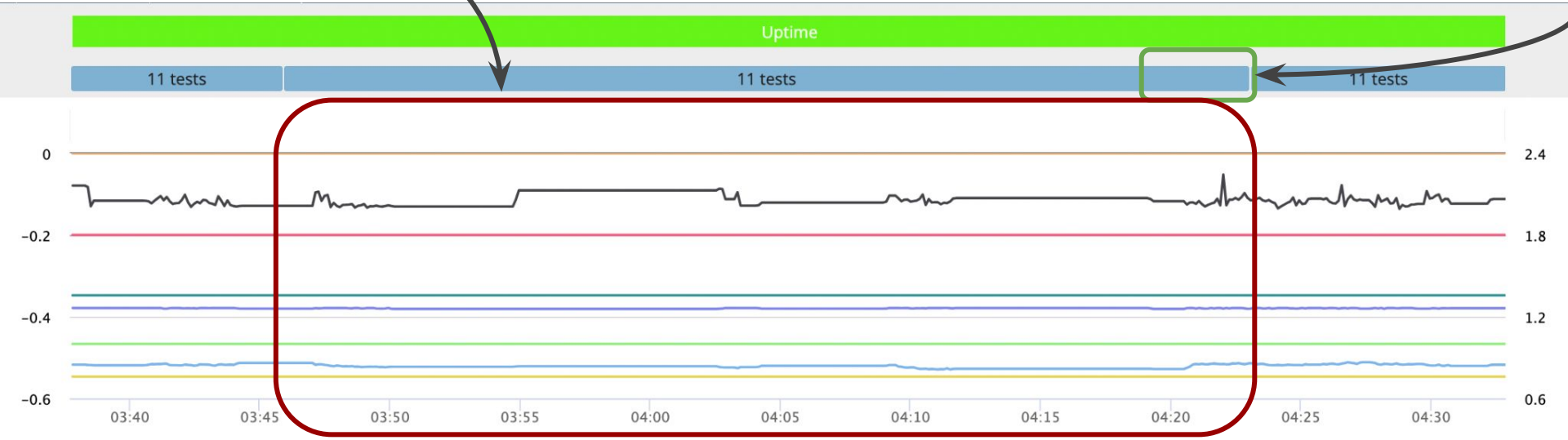
Uses: Stable Periods, Knowledge AI [in development]
Benefits: Powerful learning

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.

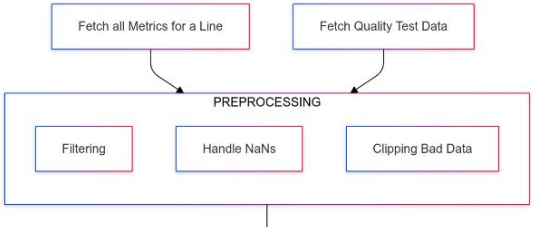
Problem Statement: Predictive Quality

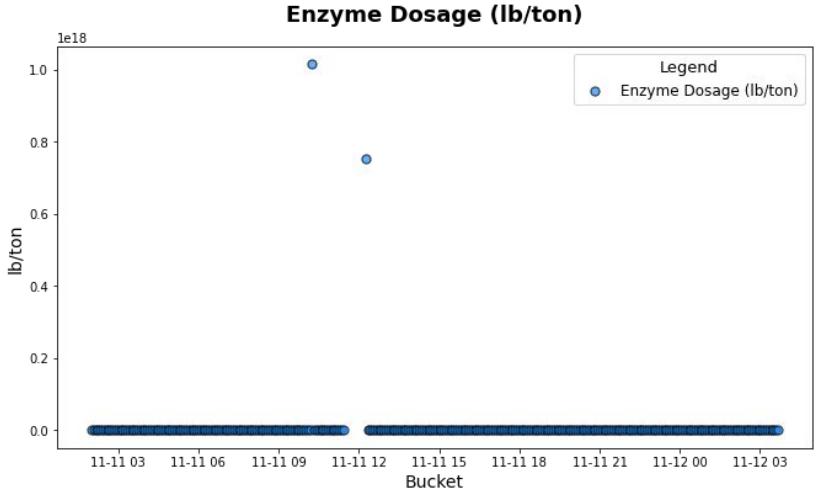
Use **multi-dimensional time series** signatures:

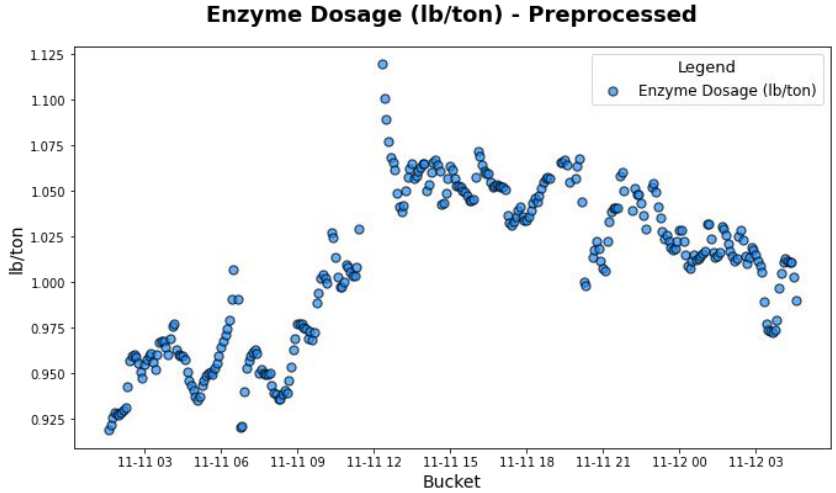
To predict a single quality test **label** at the end of each run:



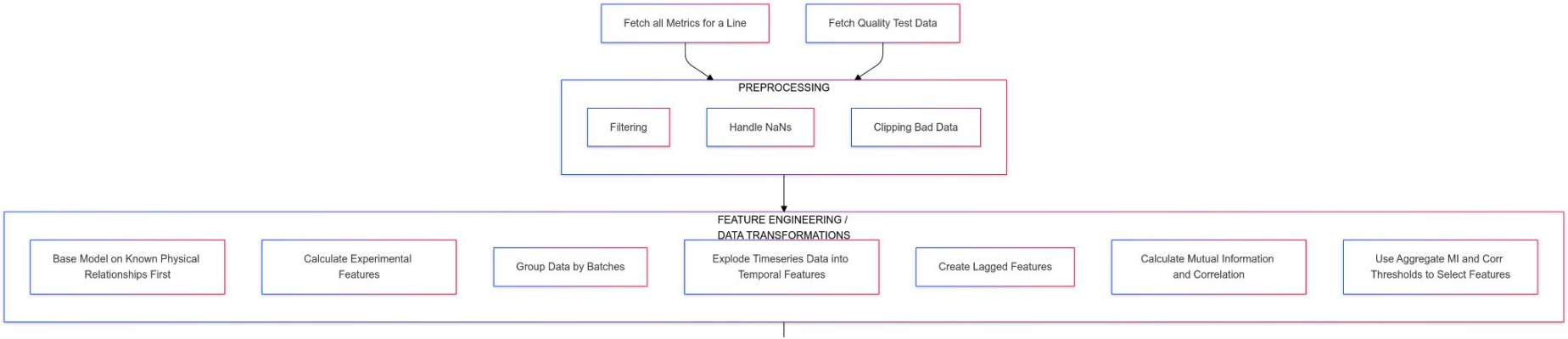
Example PQ Pipeline - Preprocessing





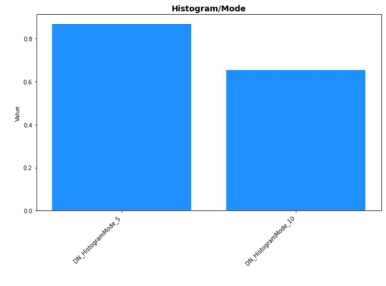


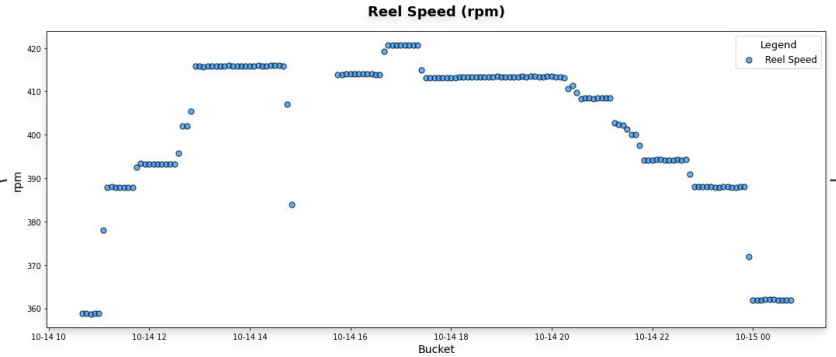
Example PQ Pipeline - Feature Engineering

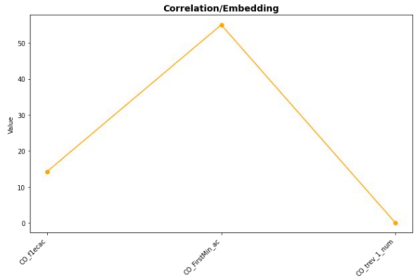


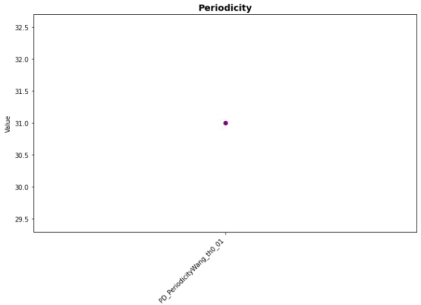
Example PQ Pipeline - Feature Engineering

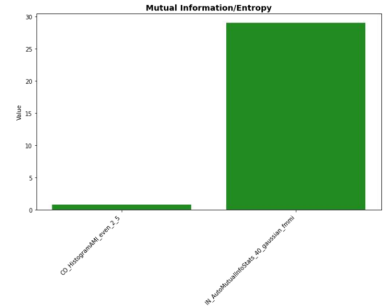
Time Series Decomposition into Core Stats

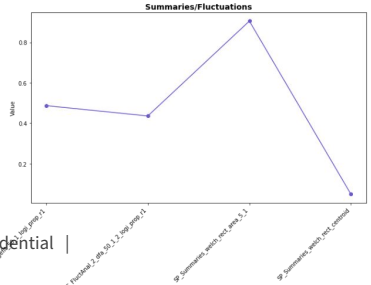


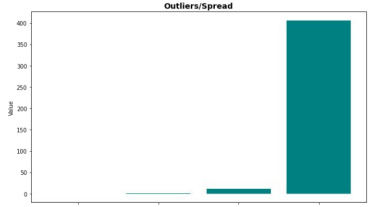


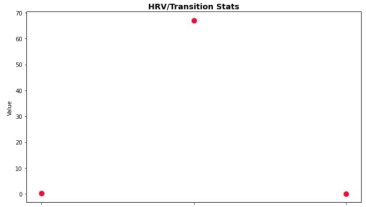












Source: catch22: CAnonical Time-series CHaracteristics
https://link.springer.com/article/10.1007/s10618-019-00647-x

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

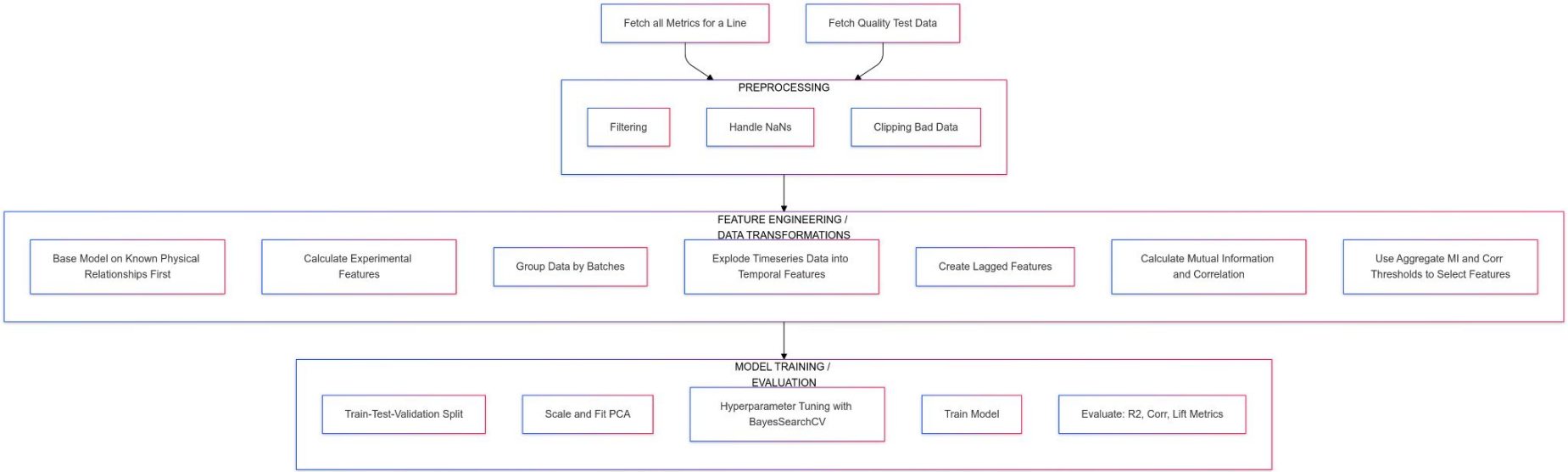
3. Concentration of Noise/"Bad Data", nunique(), etc.

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

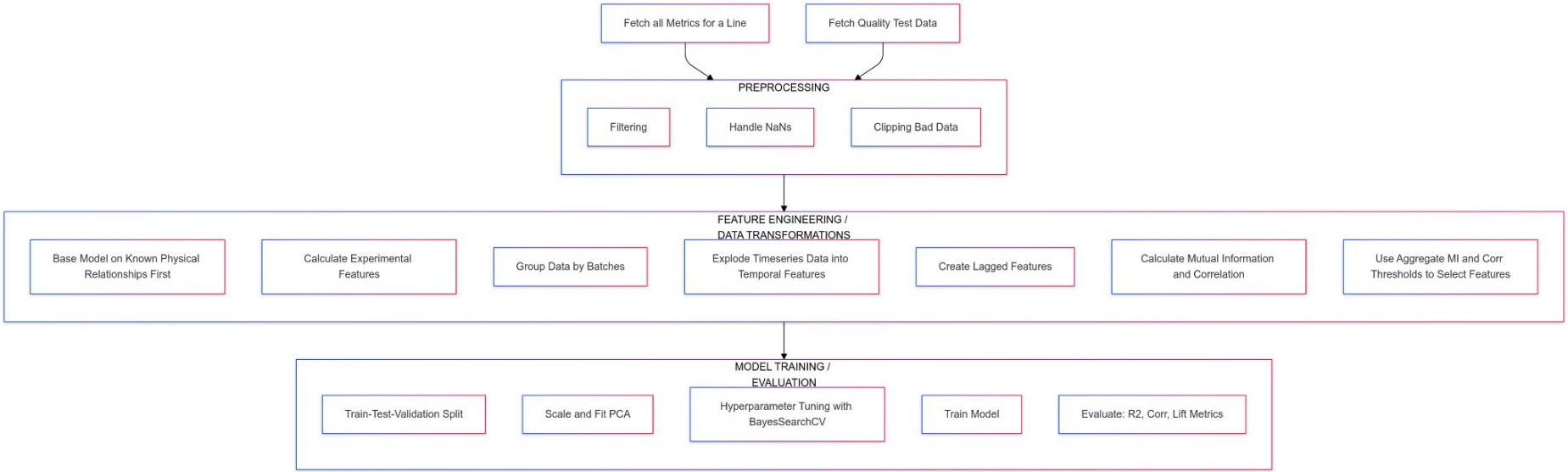
$$I(X,Y) = -\sum_{x_i, y_i} p(x_i, y_i) \log_2 \frac{p(x_i, y_i)}{p(x_i)p(y_i)}.$$

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

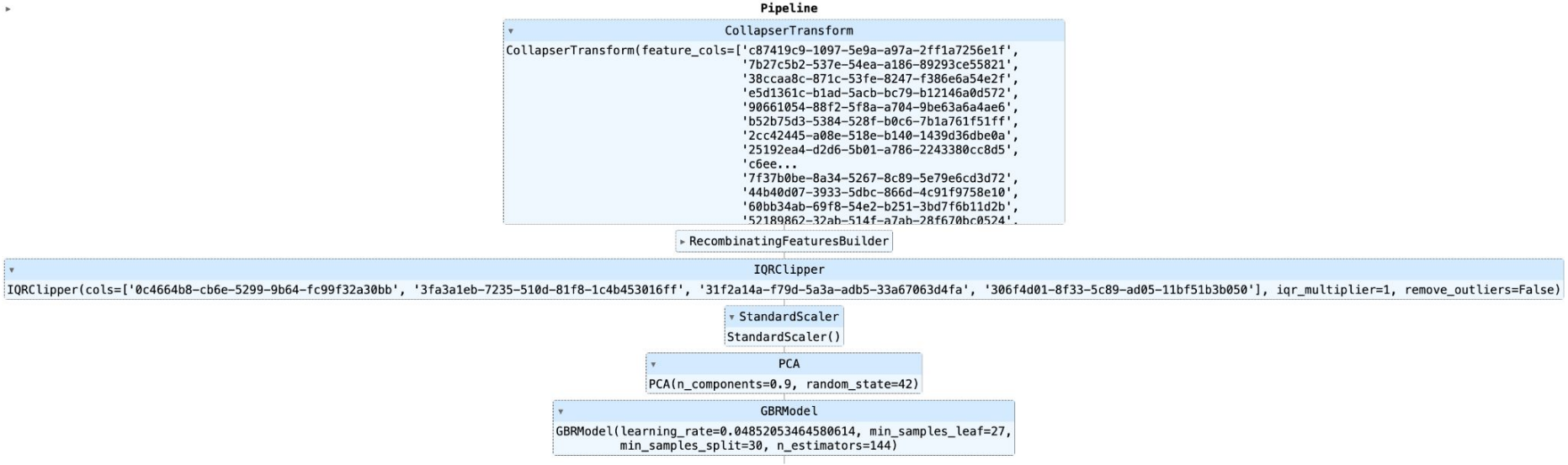
Example PQ Pipeline - Train/Deploy



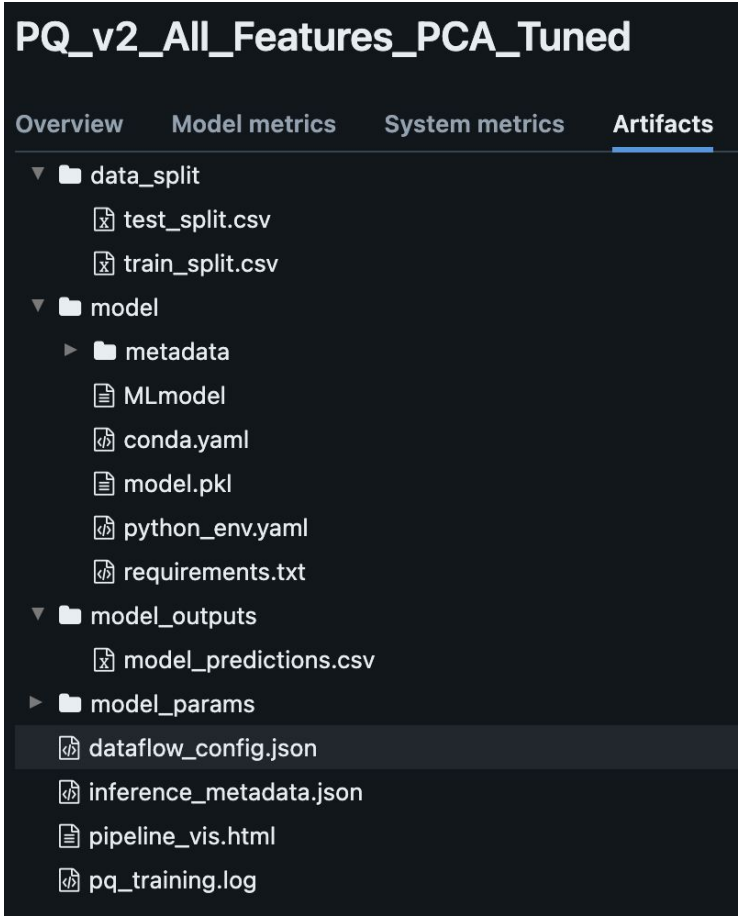
Example PQ Pipeline - Train/Deploy



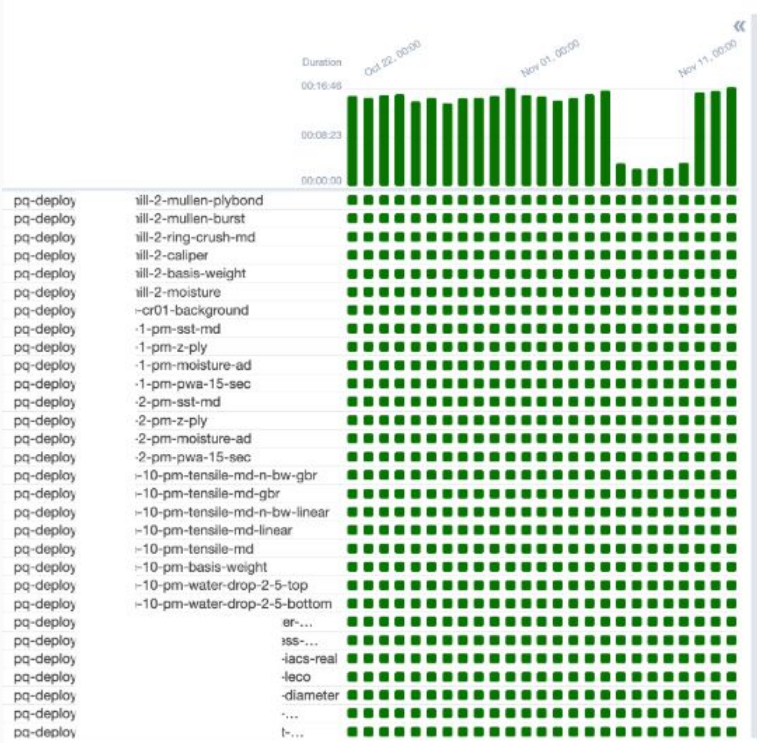
Example Pipeline - Registered to MLFlow



MLFlow Artifacts



Model Deployment via Apache Airflow



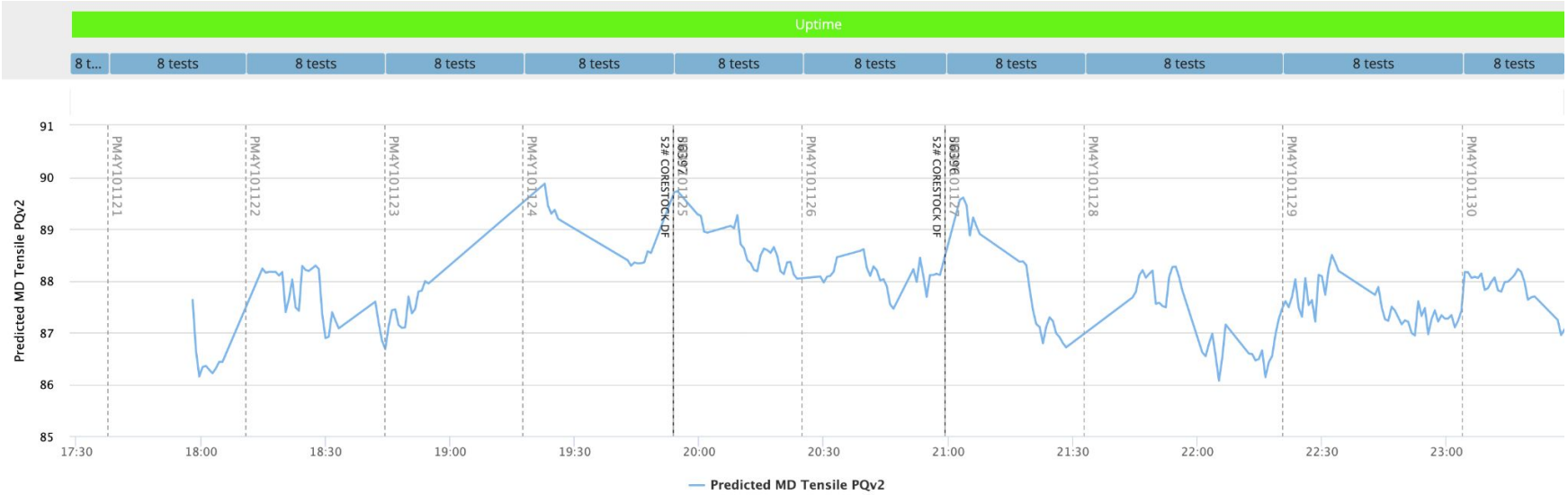
DAG
predictive_quality_deploy_v1

Details | Graph | Gantt | Code

DAG Runs Summary	
Total Runs Displayed	25
Total success	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



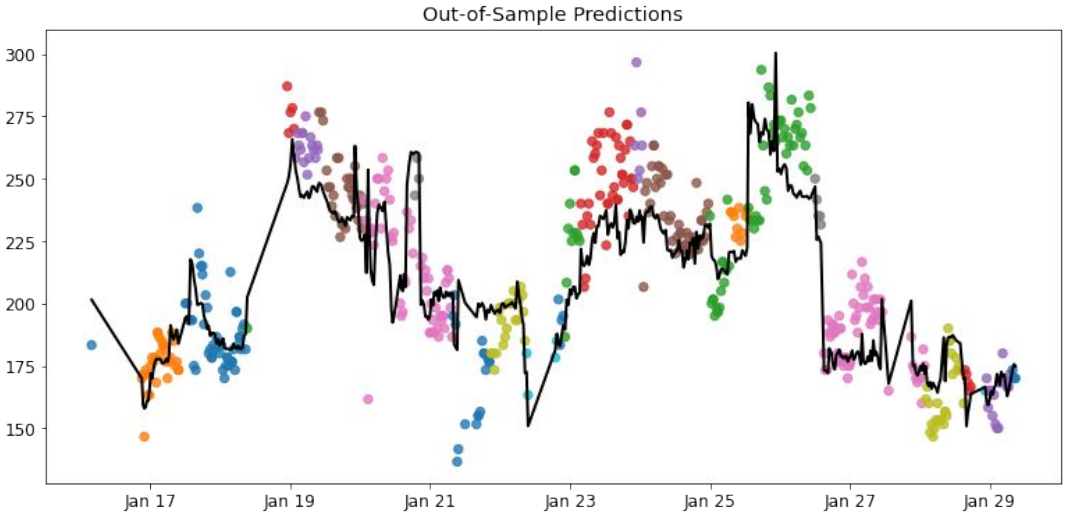
Key Variables Example: Paper

Predictive Quality Model Development

>90% accuracy

>70% correlation

<3 sigma error



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

+5.3Mft
Additional Feet Produced over 90 days

ODEN TECHNOLOGIES

Thank You!

oden.io

joseppy.ca