

Predictive Machine Maintenance Project (systemRx)

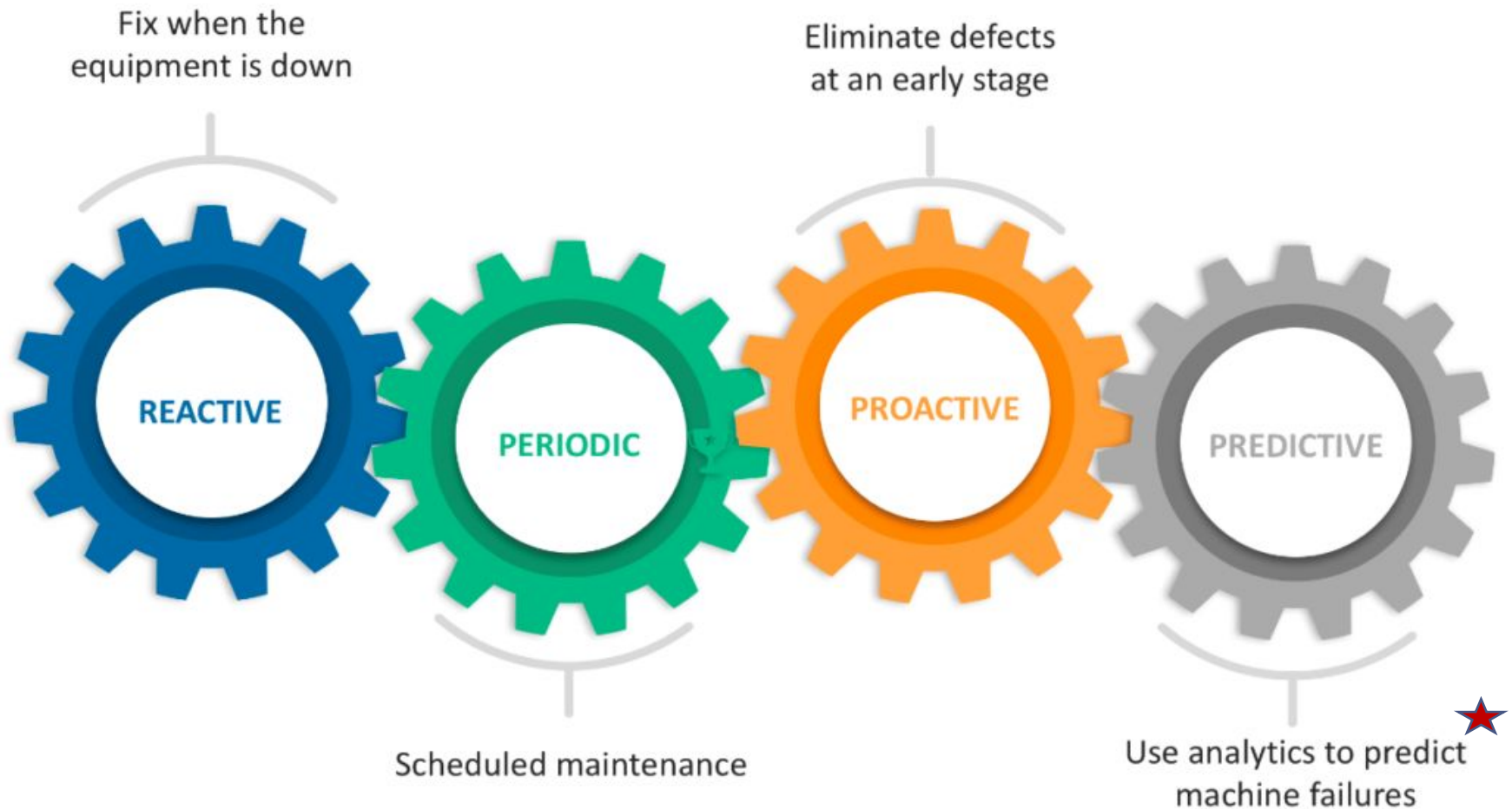


Image Source: digitaldoughnut.com

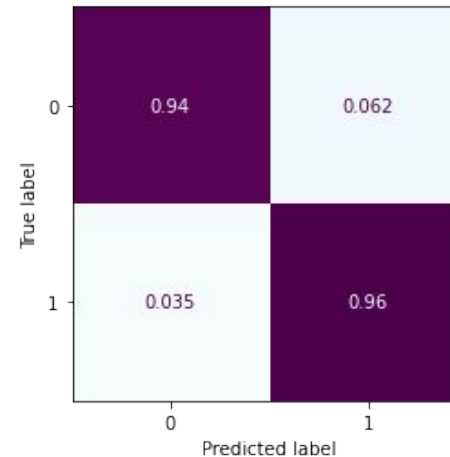
Presenter: Sheneka Allen
June 2022

Predictive Machine Maintenance Project (systemRx)

Objective

- Develop a modeling tool that accurately predicts the percent likelihood of machine failure. If possible, identify common contributors to machine failure to minimize equipment downtime and maintenance costs.
- Prediction performance goals: >90% accuracy for true positives, < 5% for false negatives
- Technology & tools: predictive maintenance (pDm) dataset (synthetic), data analytics

Visualizations



Classification Reporting for Test Set (XGB):

| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0 | 0.96 | 0.94 | 0.95 | 2405 |
| 1 | 0.94 | 0.96 | 0.95 | 2426 |
| accuracy | | | 0.95 | 4831 |
| macro avg | 0.95 | 0.95 | 0.95 | 4831 |
| weighted avg | 0.95 | 0.95 | 0.95 | 4831 |

Model 'Problem-Solving' Value

Strengths:

- Potential decrease in Mean Time Between Failure
- Potential decrease in maintenance & labor costs
- Potential reduction in single point of failure with digital monitoring / sensor data collection

Limitations:

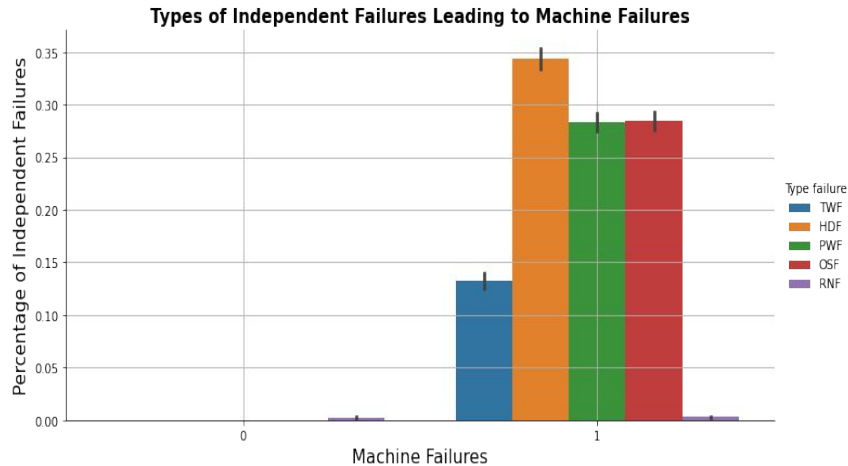
- Limited public availability for predictive maintenance data to re-train model
- Must correct imbalance datasets to ensure quality model performance (e.g., balanced dataset improved performance by **31%!!**)

Recommendations

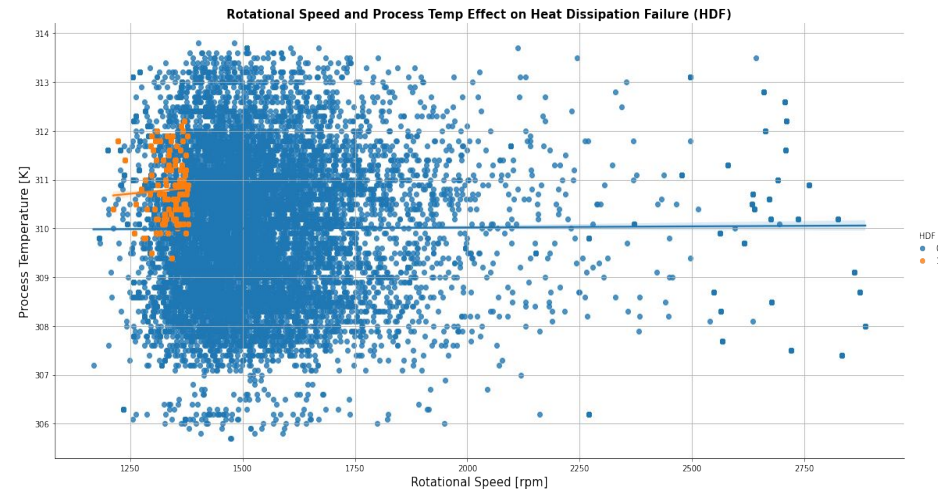
- 1) Focus initial maintenance tasks on highest percentage causes for machine failures
- 2) Develop a formal condition monitoring program
 - Identify & collect machine data on critical assets (e.g., safety and revenue loss focus)
 - Determine what typically goes wrong on machines that cause failures (e.g., 34% heat dissipation, 28% power failures)
 - Consider altering testing frequency (e.g., industry best practice: 2 data samples minimum prior to failure)

Backup - More systemRx Visualizations

Machine Failure Causes for this Dataset

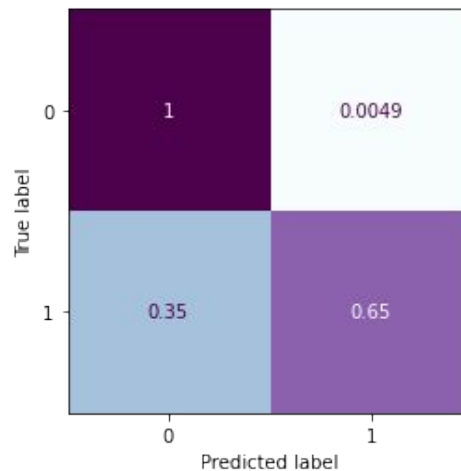
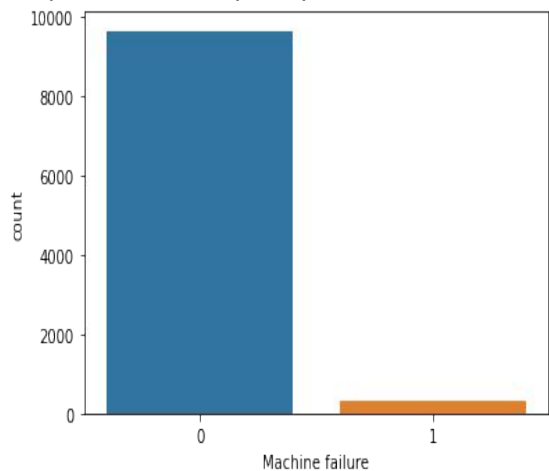


Heat Dissipation Failure - Feature Analysis



Imbalanced Data & Model Performance

Imbalanced datasets cause model predictions to be skewed in favor of the majority class versus the minority class (target). The prediction accuracy is 99% but of ZERO analysis value! You MUST correct the data imbalance and re-train the model to improve performance quality and ensure its value to stakeholders.



Classification Reporting for Test Set (XGB):

| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0 | 0.99 | 1.00 | 0.99 | 2428 |
| 1 | 0.80 | 0.65 | 0.72 | 72 |
| accuracy | | | 0.99 | 2500 |
| macro avg | 0.89 | 0.82 | 0.85 | 2500 |
| weighted avg | 0.98 | 0.99 | 0.98 | 2500 |

Presenter: Sheneka Allen
June 2022