RoboFault

Introduction & Columns Explanation

The **Universal Robots UR3** is a lightweight, compact
A 6-axis cobot designed for precision tasks such as assembly, handling, and testing.

The dataset consists of **time-series measurements**

recorded from the UR3 robot across its six joints (J0–J5), the tool

, and system-level fault indicators. The main columns include:

- **Timestamp**: Exact time of each measurement.
- **Cycle**: A logical grouping of robot operations into cycles.
- **Current_J0 … Current_J5**: Electrical current drawn by each of the

  Six joints.
- **Temperature_T0, Temperature_J1 … Temperature_J5**: Temperature sensors for the tool and each joint, indicating thermal load and heating.
- **Speed_J0 … Speed_J5**: Angular velocity of each joint, representing robot motion dynamics.
- **Tool_current**: Electrical current drawn by the tool (gripper).
- **Robot_ProtectiveStop**: Binary indicator (1 = protective stop triggered, 0 = normal operation), representing critical safety faults.
- **grip_lost**: Binary indicator (1 = gripper failed to hold the object, 0 = normal grip), representing task-level failure.

Data Type:

The dataset includes two binary fault indicators: **Robot_ProtectiveStop** and **grip_lost**. In the raw data, the **Robot_ProtectiveStop** was represented as numerical values (0 and 1). When loaded into Pandas, it appeared as float64 columns. The column was converted into the Boolean data type.

Handling the Nulls:

The data has a small percentage of nulls, but as its log is taken each second and its per cycle, so I think that dropping the nulls isn't the best solution, as if the cycle is small and 10 seconds, and I have 5 seconds of them as nulls, and i removed them I by that will ruin the way the cycle was working.
So we have experienced 3 methods and compared them to our data.

1. **Linear Interpolation (per cycle)** – simple time-based interpolation within each cycle.

2. **KNN Imputation** – imputes based on similarity to nearest neighbors.

3. **MissForest-like Imputation** – a machine-learning-based approach using iterative imputation with ExtraTrees.

Linear Interpolation

- Works cycle by cycle.

- Estimates missing points by connecting the nearest valid values in time.

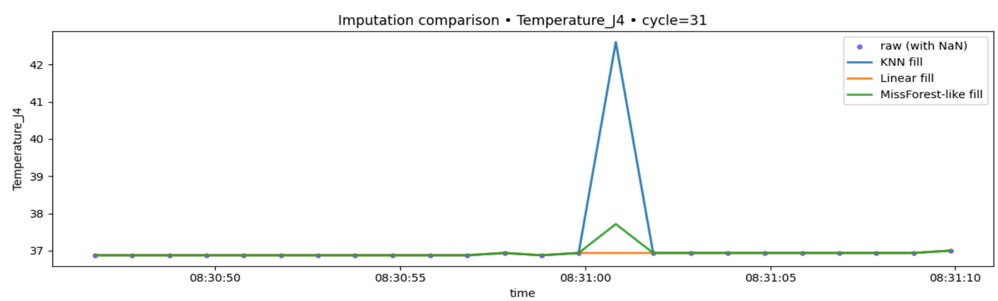
- **Limitation:** if nulls are at the start/end of a cycle, interpolation has no neighbors and can propagate bias.

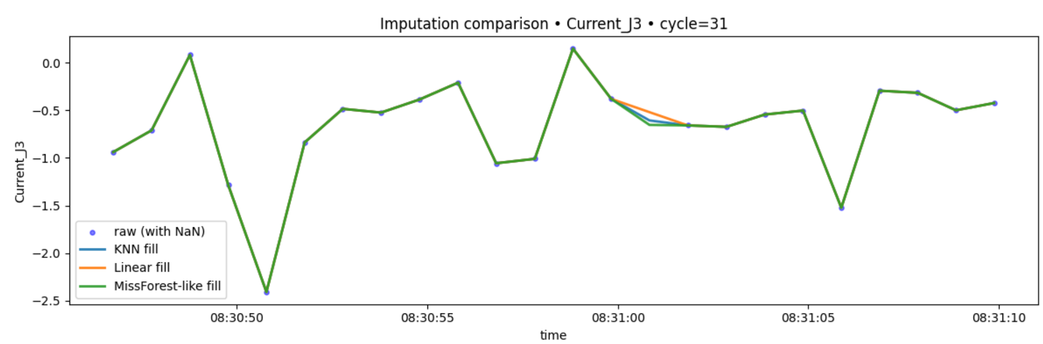
KNN Imputation

- For each missing value, find the *k* nearest complete samples (in feature space) and fill using their average.
- Handles cross-feature dependencies.
- **Limitation:** can oversmooth sudden spikes in our data that appear in the temperature (affected by outliers).

MissForest (IterativeImputer + ExtraTrees)

- Treats imputation as a supervised learning task: predicts each missing feature from all others.
- Iteratively refines values until convergence.
- **Strengths:** preserves complex nonlinear dependencies, robust to outliers.





At the end, we choose to go with the linear interpolation as it’s:
As it's the simplest way and the least complex that serves the data as a timeseries, and the nulls are small percentages, so the linear interpolation is what we chose

And for the **Robot_ProtectiveStop** nulls, which are categorical (binary: stop or not stop), we imputed missing values using the **mode within each cycle**. Specifically, for every cycle, we checked the most frequent value (0 or 1).

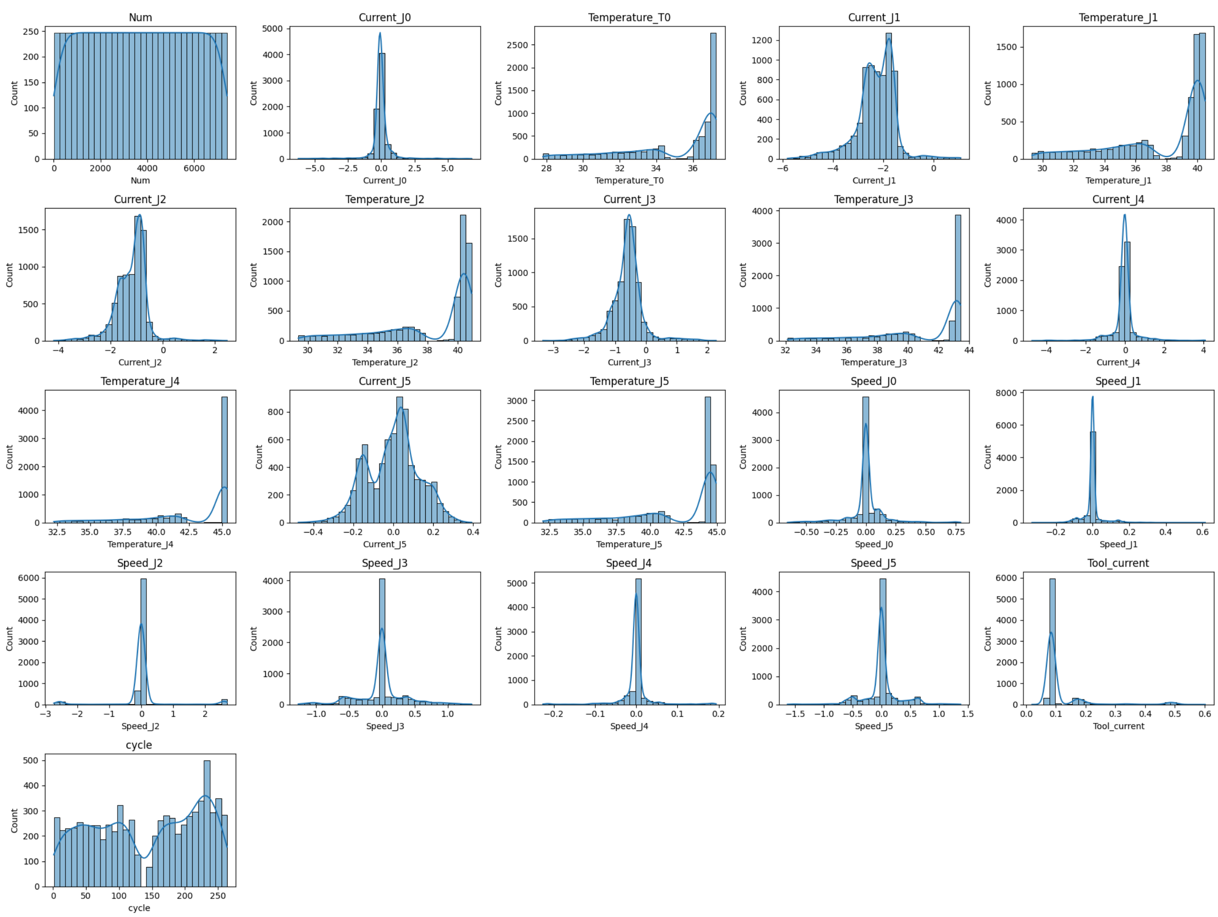
Data Distribution & Correlations:

Joint Currents: Most joints (J0–J5) follow approximately centered distributions, but with some skewness and long tails. This indicates occasional **spikes in current demand**, possibly linked to faults or heavy loads.

Joint Speeds: Speeds are sharply peaked around zero, with many small values and a few strong outliers. This suggests the robot mostly operates with small corrective movements, with occasional bursts.

Temperatures: All joint temperatures (J1–J5) increase steadily, showing narrow peaks with some long upper tails. Outliers here may correspond to overheating events.

Tool Current: Tool current shows a very narrow distribution with small deviations, but with some significant outliers during faults.

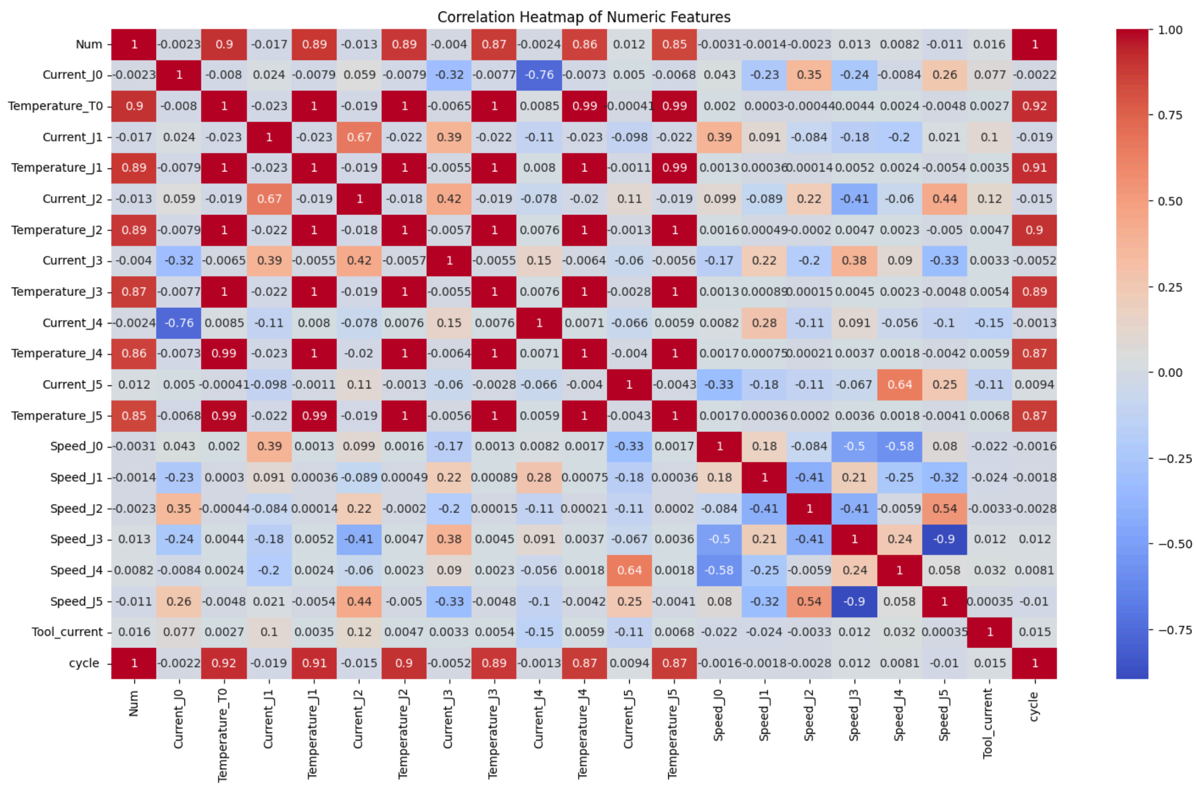


Strong positive correlations are found between the **temperature sensors** (J1–J5 and T0). This makes sense, as heat builds up across the entire system simultaneously.

Moderate correlations exist between some **currents and speeds**, e.g., Current_J1 with Speed_J1..

Negative correlations are visible, for example, between Current_J0 and Temperature_J4. These may highlight compensatory behavior: one joint working harder while others stabilize.

Tool Current shows **weak overall correlations**, suggesting it is more **independent** from joint-level behavior.



The Strong Correlations between the Temperatures and how they act like each other can only be a statistical measure for them in the machine learning model, not all of them.
Also, the Protective_stop and grip loss are unbalanced, which can indicate that the faults are a small percentage : 0: 96.25%,1: 3.75%

Feature Engineering

1. Absolute Values

- **Features**: abs_Current_J, abs_Speed_J

- **Why**: Raw currents and speeds can be positive (movement in one direction) or negative (opposite direction). For fault detection, the **magnitude** is often more important than the sign.

2. Power Features

- **Features**: power_J = Current_J × Speed_J

- To see what happens or what the faults are associated with the increase in power

3. Sudden Change Features

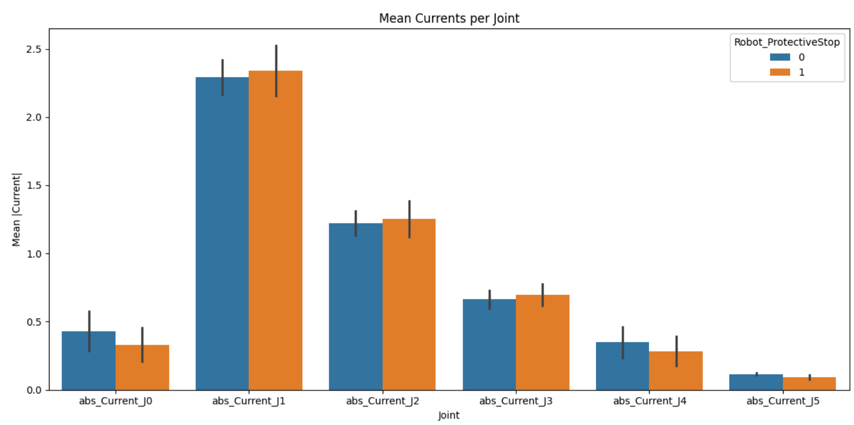
- **Features**: sudden(Acceleration)

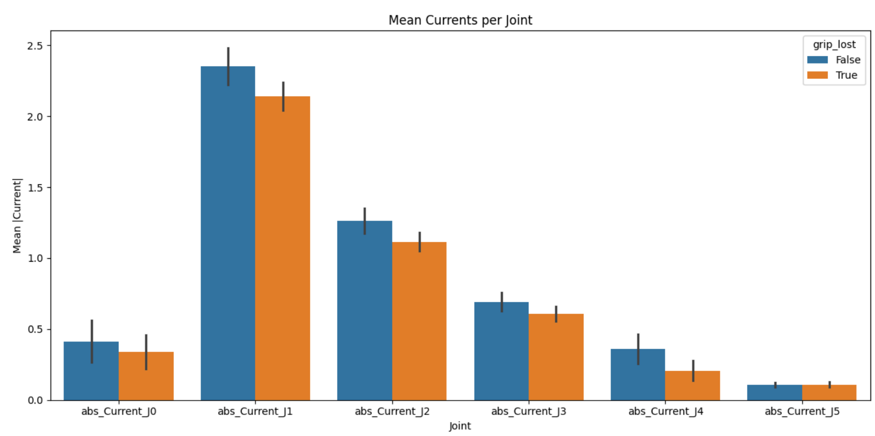
- **Why**: Rapid changes in current or speed can lead to Faults

Insights From The new Features:

Absolute Joint Currents vs. Protective Stop / Grip Lost

- J1 and J2 have **much higher mean currents** than other joints, confirming they bear most of the robot’s load.
- Protective stops often occur when **currents spike slightly above normal**, especially for J1 and J2.
- Grip loss correlates with **reduced current** in some joints (J2, J3), suggesting that loss of grip is not affected by the increase in current

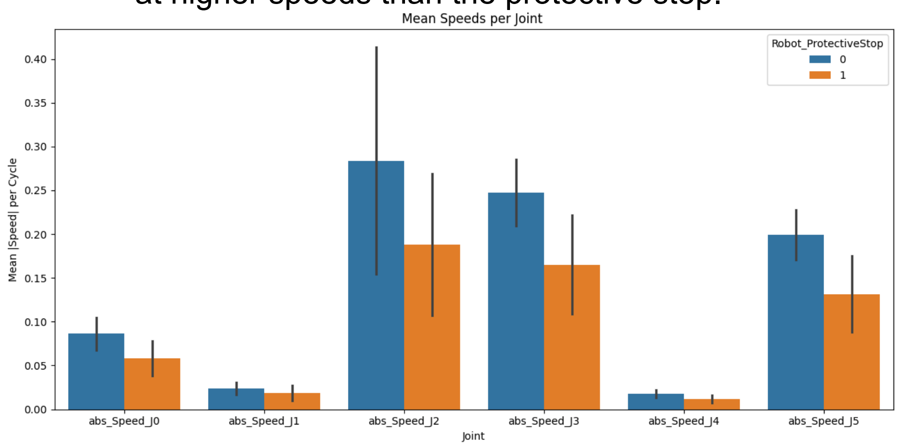


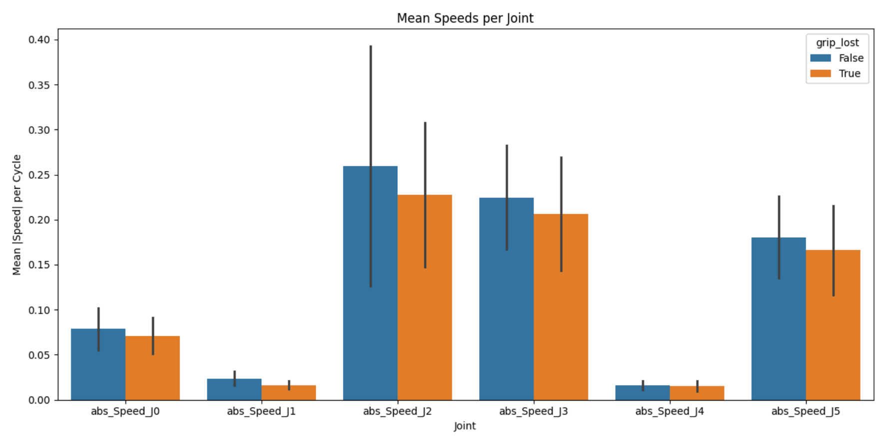


Absolute Speeds vs. Protective Stop / Grip Lost

- Speed is **consistently lower during protective stops**, across almost all joints. Suggests stops are triggered **when motion slows abnormally**

- For grip loss, speed is associated with the grip_lost more so we can see the fault at higher speeds than the protective stop.



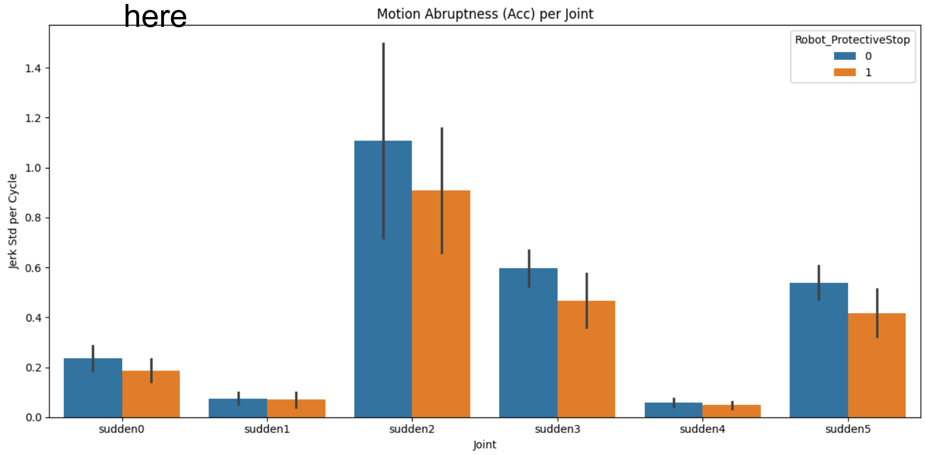


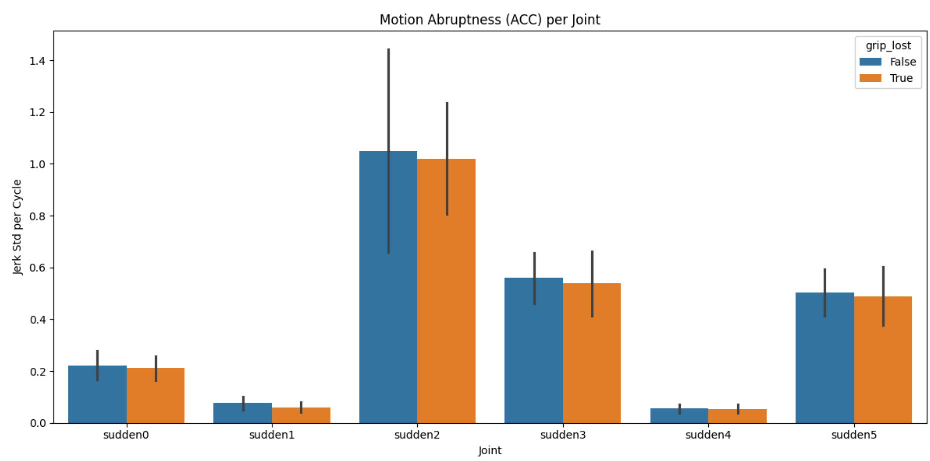
3. Motion Abruptness

- J2 and J3 show the largest **variability in Motion Abruptness** .

- During protective stops, the sudden change doesn't affect it like in the grip loss, where the stops don't happen as a result of the sudden change in the speed or the current.

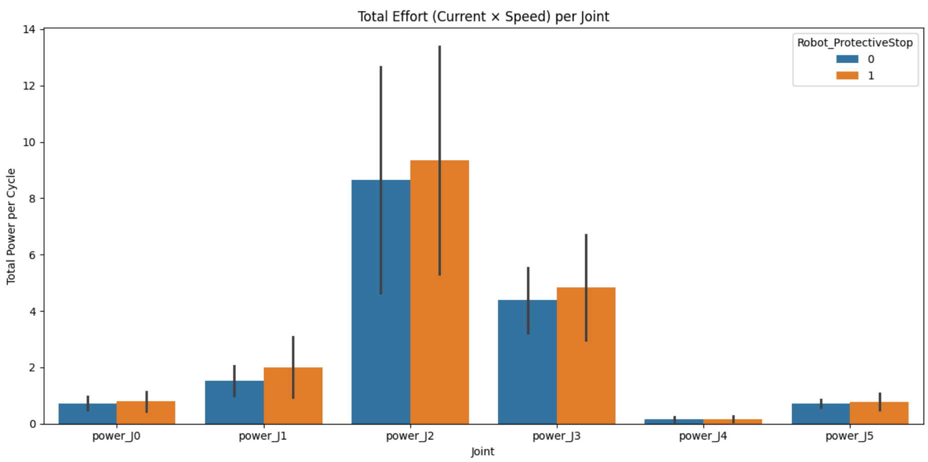
- During grip loss, sudden(acc), the sudden affects here more than that at the protective stops. We see that the grip loss is more true in the sudden change here

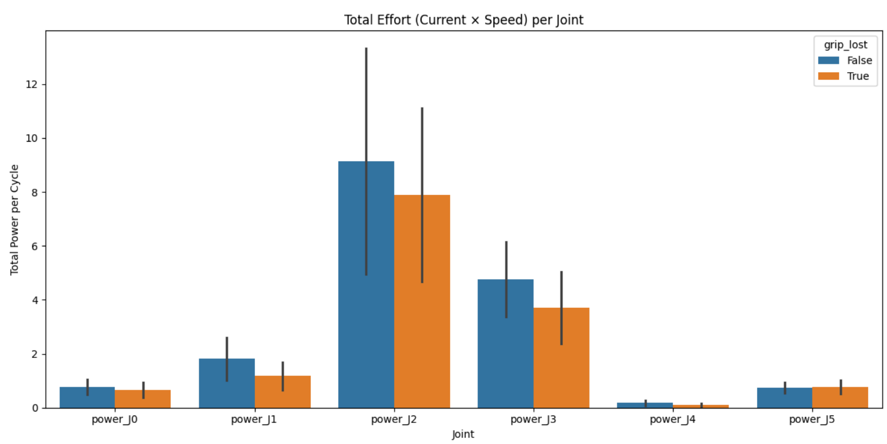




Power per Joint

- J2 dominates in total power consumption, followed by J1 and J3's strongest load.
- Protective stops show **slightly elevated power in J3 & J2**, indicating overload conditions.
- Grip loss shows **lower power consumption across joints,** indicating that the increase in the current is not the thing that affects the grip loss.





Current Asymmetry (Forward vs Backward)

The current asymmetry:

- Measures if the robot uses **more current in one direction than the other**.

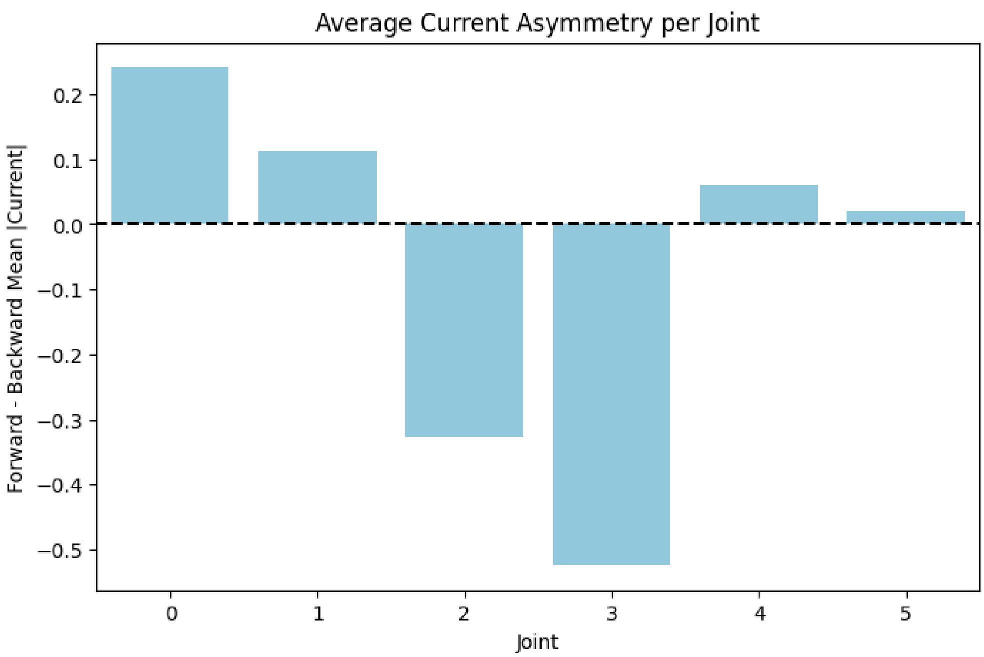
How it's calculated:

- Split movements into forward vs backward cycles.
- Compute the mean absolute current for each direction.
- Subtract: (Forward mean − Backward mean).

Interpretation:

- **Near-zero** Symmetric behavior.
- **Positive/Negative bias** means it needs more current in a specific direction.

- Joints J2 and J3 showing strong negative asymmetry means they struggle more in backward motion.



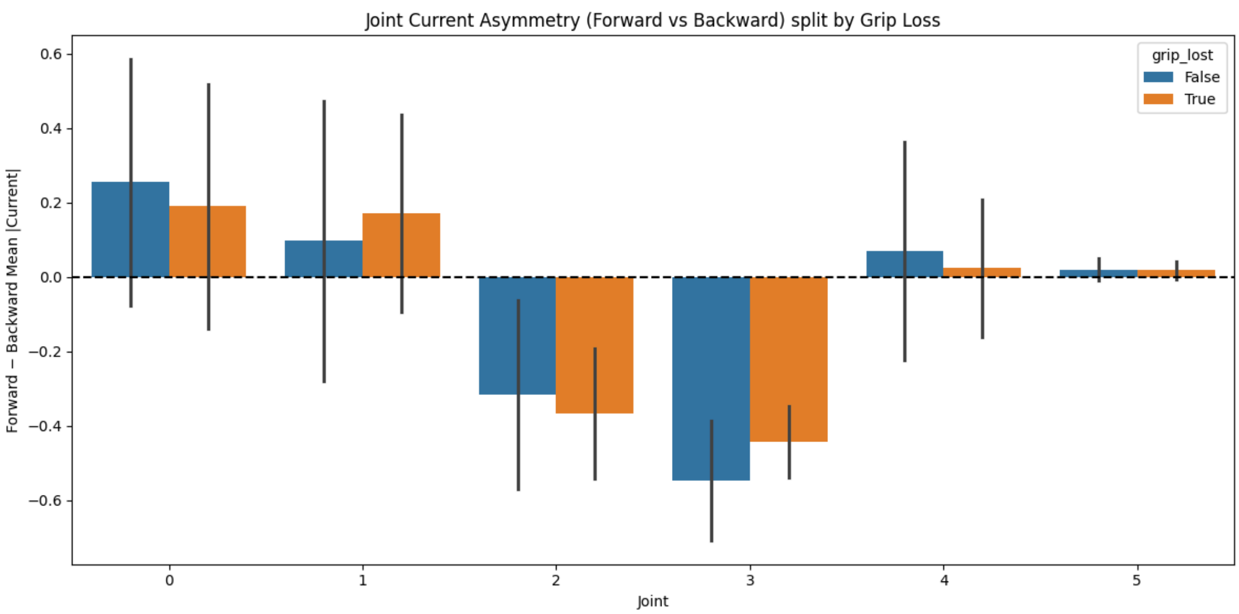
Current Asymmetry split by Grip Loss

What it is:

- Same forward-backward asymmetry, but now grouped by whether the robot experienced grip loss.

Interpretation:

- If asymmetry **worsens during grip loss**, it means unstable gripping affects mechanical balance.
- J2 and J3 show larger negative asymmetry under grip loss confirming mechanical strain.



Temperature vs Power Correlation

What it is:

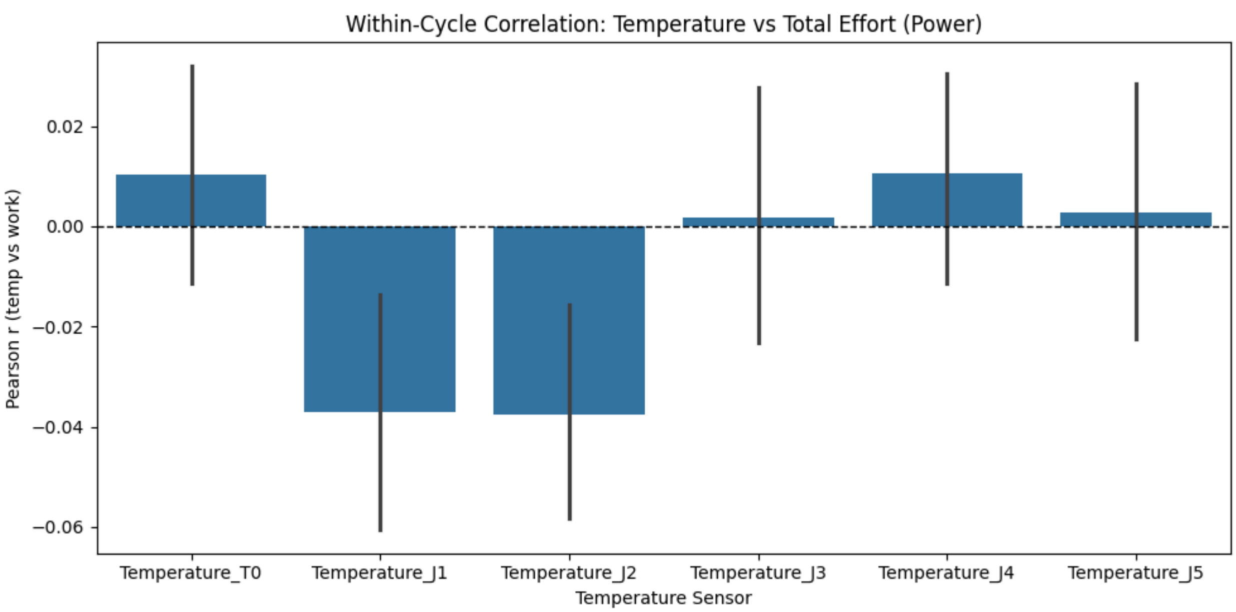
- Tests whether joint temperature correlates with mechanical work (power = current × speed).

How it's calculated:

- Within each cycle, compute the Pearson correlation between **temperature rise** and **power exerted**.

Interpretation:

- **Strong correlation**: Heat is proportional to workload (healthy thermodynamics).
- **Weak/no correlation** Either cooling issues, sensor errors, or thermal runaway.
- **weak correlation in some joints**, meaning heat is not always explained by workload as a risk factor or lag in the sensors.



Thermal Lag

What it is:

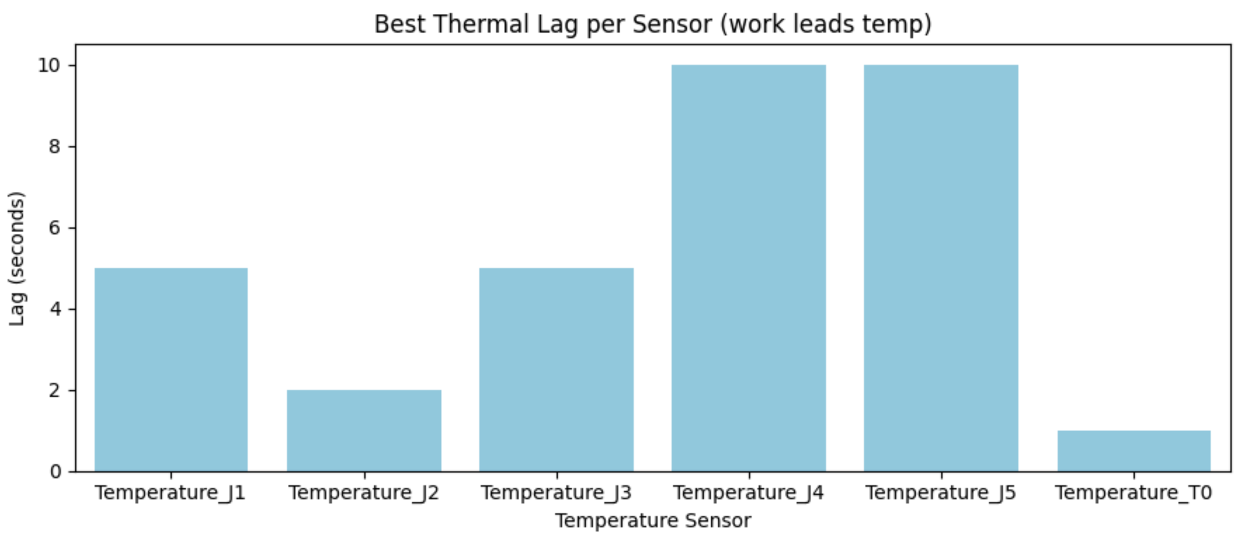
- Checks how long after applying work (power), the temperature rises.

How it's calculated:

- Cross-correlation between power and temperature signals.
- Lag (in seconds) where correlation peaks = thermal response delay.

Interpretation:

- **Short lag (2–5s)** Normal heating.
- **Long lag (10s+),** Slow thermal response, possibly due to cooling inefficiency or sensor lag.
- Joints J4/J5 showa **large lag (~10s)**.



Peak Temp–Work Correlation at Best Lag

What it is:

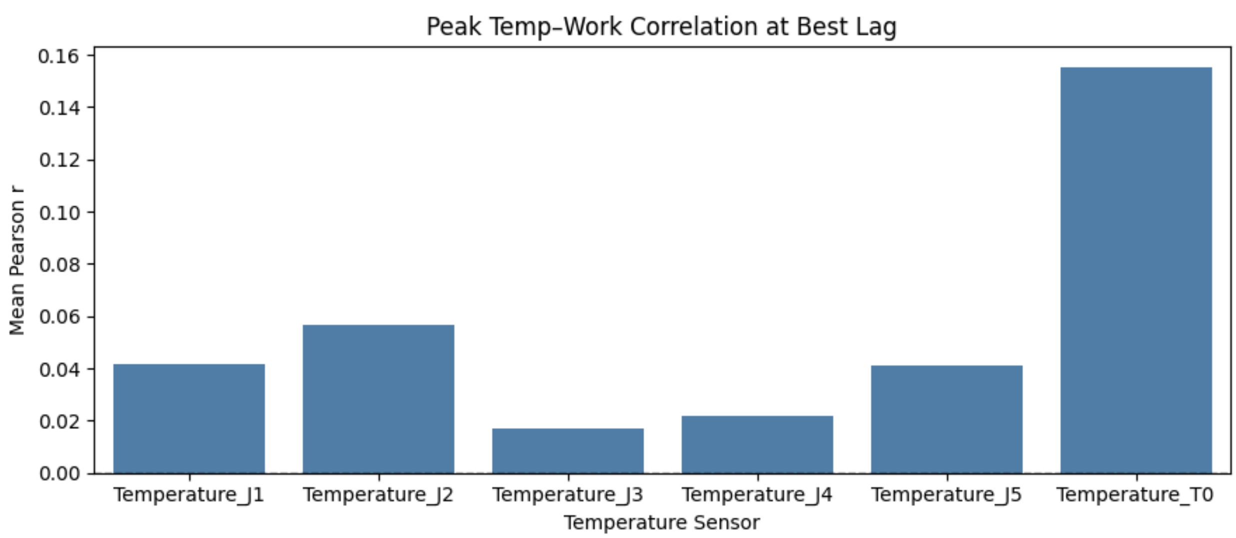
- Same as above, but now we measure the *strength* of correlation at the optimal lag.

How it's calculated:

- Take Pearson r at the lag with maximum correlation.

Interpretation:

- Joints with **higher correlation** (like T0 and J2) show predictable heating.
- Joints with **lower correlation** may have cooling problems or faulty sensors.



1. Protective Stop Events

- We filtered rows where Robot_ProtectiveStop == 1.
- Events were grouped if they occurred within **1 second apart**, so each group represented a distinct stop.
- For each group, we calculated:
  - **Start and End time**
  - **Duration**
  - **Average Joint 0 Speed, Current, and Temperature**

Findings (Top 5 Protective Stops):

- Most stops lasted **5–11 seconds**, showing they were **short disruptions** rather than long failures.
- Average temperatures were **~32–37°C**, which are not critically high, so overheating was **not the main cause**.
- Average currents were **near 0** (some negative, some slightly positive), suggesting the robot was **not under heavy load**.

Non-Protective Periods (Normal Operation)

- We repeated the same grouping for Robot_ProtectiveStop == 0.
- The **longest continuous runs** lasted **200–338 seconds** without interruptions.
- These normal operations showed stable currents (~−0.05 to +0.01) and expected temperatures (~28–36°C).
- Compared with stops, we see that **Protective Stops break continuous cycles**, but **do not necessarily coincide with extreme sensor values,** pointing to **some other features** rather than overheating.

Grip Lost Events

- Similar grouping was done for grip_lost == 1.
- Grip loss events were typically **short (9–18 seconds)**.
- During grip loss:
  - Currents were again small, with one spike at **0.6 A**, showing the robot **increased current to maintain grip** but failed.
  - Temperatures were in the **28–37°C** range, again not overheating-driven.

Non-Grip Lost Periods

- Long continuous segments (up to **963 seconds**) occurred without grip loss.
- Currents stayed negative or near zero, with stable moderate temperatures.
- This contrast shows that **grip losses are short-lived anomalies**, while normal cycles are long and stable

Overall Insight:

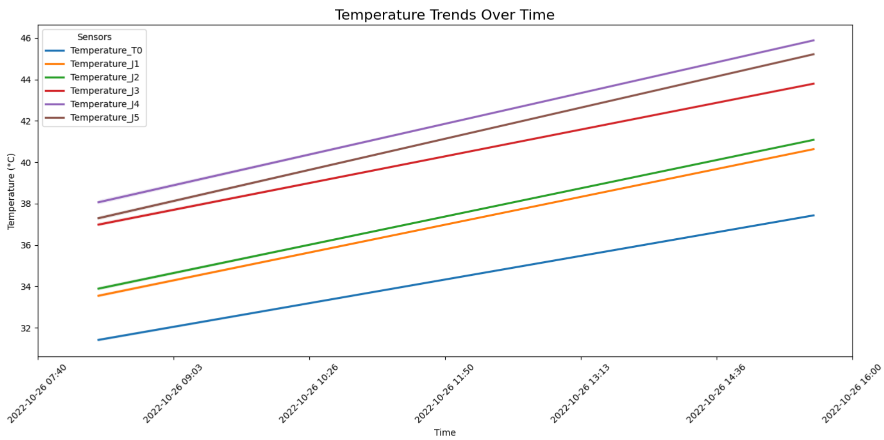
- **Protective Stops** and **Grip Loss** events are **short-lived disruptions** compared to long stable periods of normal operation.

Temperature and Tool Current Analysis

Temperature Trends

- Across all sensors, temperature increased steadily over time, reflecting the expected **thermal accumulation effect** from continuous operation.

- Higher-temperature joints (J4, J5) showed a slightly faster rise compared to cooler joints (T0), suggesting different **thermal loads**.



Tool Current Behavior

1. **Mean & Max Tool Current with Faults**

   - Tool's current shows sharp fluctuations during protective stops and grip loss events.

   - Protective stops are more often associated with **higher tool current spikes**, while grip loss appears when current fluctuates around the normal range.

2. **Tool Current Variability**

   - The standard deviation of tool current is a strong indicator of instability.

   - Fault events (both stops and grip losses) are concentrated in cycles where variability is high, suggesting that **instability is as critical as load level**.

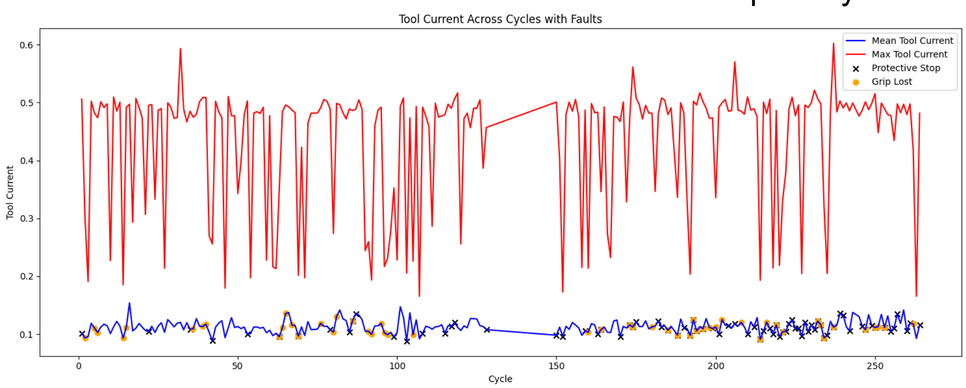
3. **Rolling Trend of Tool Current (20-cycle window)**

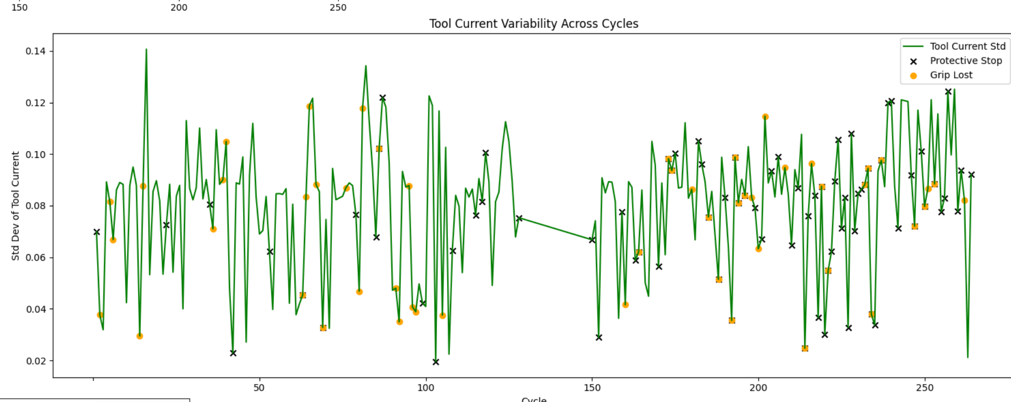
   ○ The smoothed rolling average reveals slow **drifts in tool current** that are less visible in the raw data.

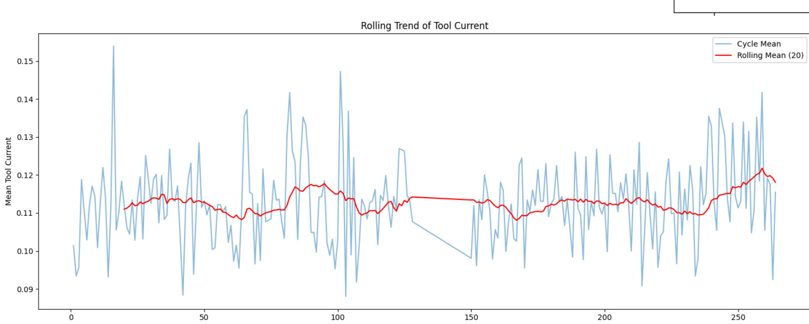
4. **Tool Current vs Fault Rate**

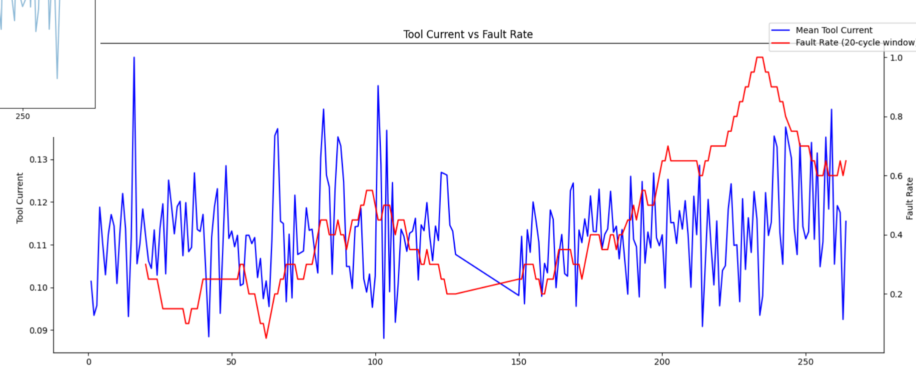
   ○ When overlaid with fault rate, the analysis shows a clear positive relationship:
   A **higher mean tool current correlates with increased fault probability**.

   ○ This confirms that tool load is a primary driver of protective mechanisms.





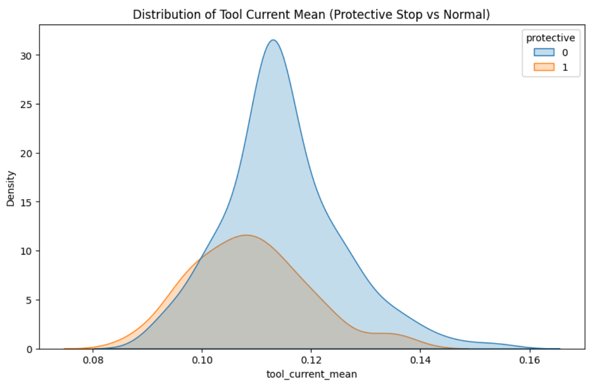


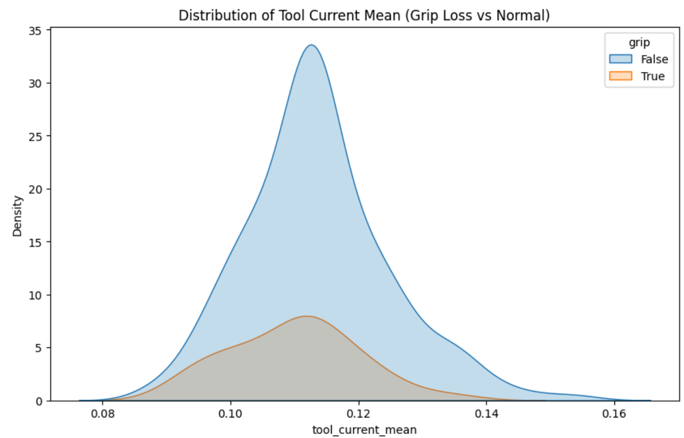


Distributional Analysis

- **Protective Stop vs Normal**: The distribution shifts toward slightly lower average tool currents during protective stops, but with a wider spread, suggesting that **instability rather than absolute level** triggers stops.

- **Grip Loss vs Normal**: Grip loss events are also associated with a lower and more dispersed current distribution, indicating **inconsistent gripping forces**.





Joint-Level Currents

1. **Load Differences**

    - Joint **J1** consistently draws the **highest average current**, followed by **J2**.

    - This suggests that these joints bear the **largest mechanical load** during the robot's operation (likely because of their role in supporting and moving heavier parts of the arm).

    - In contrast, joints like **J0, J4, and J5** consume much lower current.

2. **Fault Associations**

    ○ **Protective stops** frequently coincide with elevated or unstable current levels in **J1 and J2**, confirming that these joints are critical stress points.

    ○ **Grip loss events** also tend to appear in cycles where **J1 and J2 currents are unstable**, reinforcing the idea that load imbalances contribute to loss of grip stability.

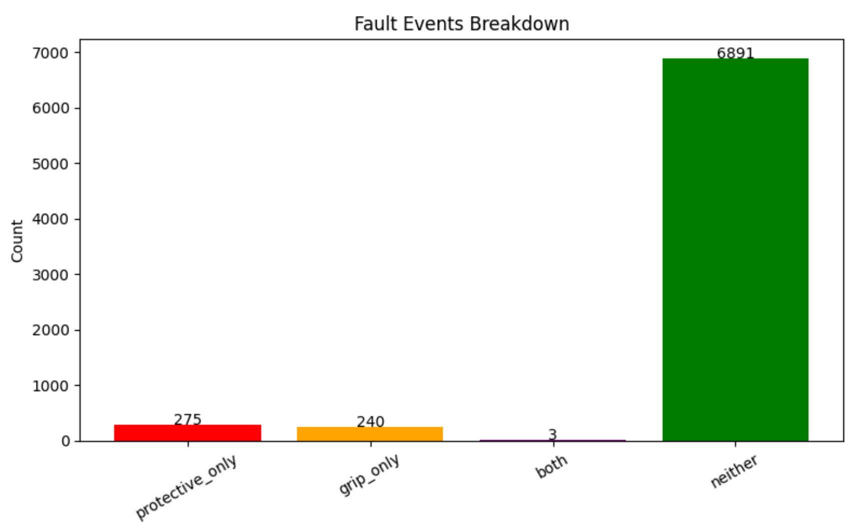
3. **Imbalanced Load as a Risk Factor**

    ○ Because J1 and J2 handle higher currents.

    ○ When these joints experience spikes or fluctuations, the protective systems are triggered to avoid overheating or mechanical damage.

    ○ This **unequal distribution of effort** across the joints is a potential **root cause for fault events**: even if the total system current is within range, localized overload in J1/J2 destabilizes the robot.

<u>The Number of faults and if they happen together:</u>



The protective stop is the one that frequently happens, and both happen only three times.

<u>Outlier and Fault Detection Analysis</u>

Boxplot Distributions

Boxplots of joint currents, joint speeds, joint temperatures, and tool current highlighted clear imbalances across the system.

- **Joint Currents:** J1 and J2 consistently carried higher baseline loads and exhibited wider variability, indicating disproportionate stress.

- **Joint Speeds:** Most joints showed narrow speed distributions centered around zero, but J2 displayed larger fluctuations, suggesting irregular motion.

- **Temperatures:** All sensors followed an upward trend, with J4 and J5 operating at the highest ranges.

- **Tool Current:** Generally stable but prone to sudden spikes, especially during grip loss and protective stop events.

Outlier Counts Across Cycles

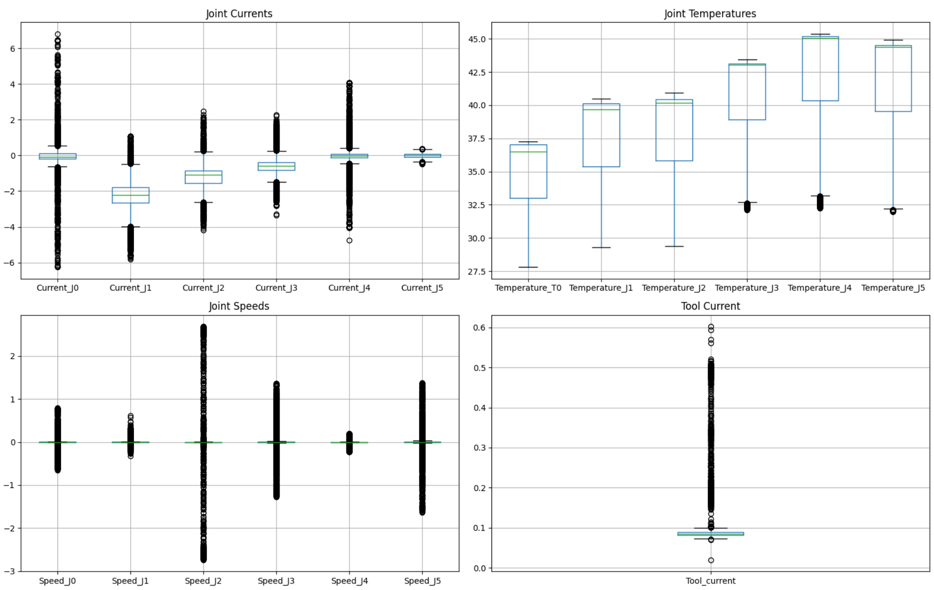
By tracking the number of statistical outliers per cycle, we observed that cycles with protective stops or grip losses showed a **sharp increase in outlier density**. In some cases, fault cycles exceeded **200 outliers per cycle**, compared to ~80–100 during normal operation. This demonstrates that outlier frequency can serve as an early warning signal for faults.

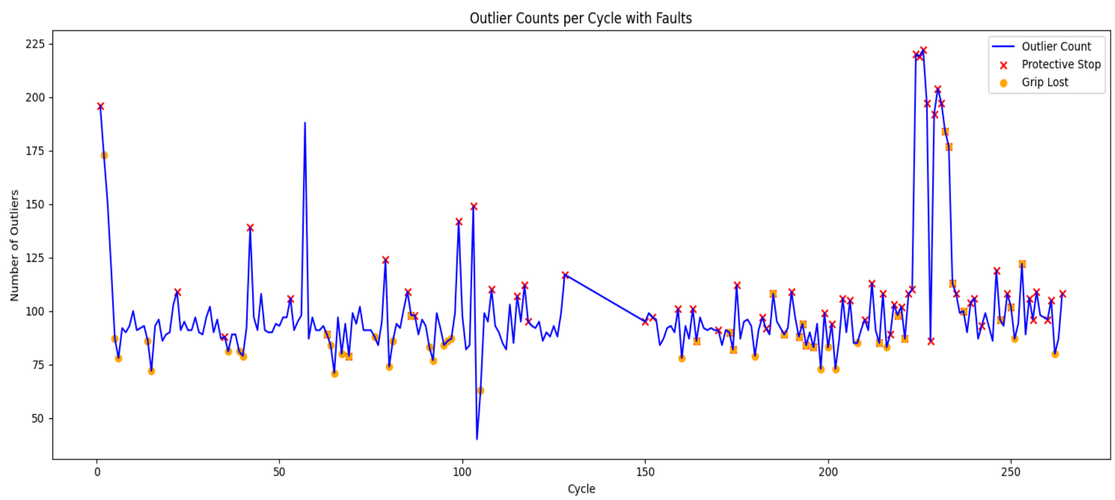
When examining feature contributions, **joint speeds were the largest source of outliers**, particularly at joints J3, J0, and J5. Currents (J4, J0, J3) and tool current also contributed significantly, reinforcing the view that faults emerge from a combination of load imbalance and motion irregularities.

Key Insights

1. **Speed anomalies are the most consistent fault indicator**, dominating the outlier counts during fault cycles.

2. **Grip loss detection is more reliable than protective stop detection** using simple outlier methods, as shown by higher recall values.


Joint Currents
Current_J0
Current_J1
Current_J2
Current_J3
Current_J4
Current_J5
Joint Temperatures
Temperature_T0
Temperature_J1
Temperature_J2
Temperature_J3
Temperature_J4
Temperature_J5
Joint Speeds
Speed_J0
Speed_J1
Speed_J2
Speed_J3
Speed_J4
Speed_J5
Tool Current
Tool_current


Outlier Counts per Cycle with Faults
Outlier Count
Protective Stop
Grip Lost
Number of Outliers
Cycle

Overall Summary of Insights

Our analysis of the UR3 robot dataset revealed several key insights into the conditions leading to protective stops and grip loss events:

1. **Imputation & Data Quality**
    - Missing values were best handled with a **MissForest imputation**, which preserved nonlinear dependencies between sensors (currents, speeds, temperatures, and tool current).
    - This ensured realistic signal reconstruction and avoided biases that could distort fault analysis.
2. **System Load & Joint Behavior**
    - **Joints J1 and J2 consistently carried the highest electrical currents**, indicating they bear most of the mechanical load.
    - Faults (both protective stops and grip losses) were strongly associated with spikes or instability in these joints, highlighting them as **critical stress points**.
3. **Temperatures & Thermal Dynamics**
    - All joint temperatures rose steadily during operation, with **J4 and J5 reaching the highest levels**.
    - Thermal lag analysis showed **delayed heat response** (up to ~10s) in some joints, suggesting cooling inefficiency or sensor lag.
4. **Tool Current as a Fault Indicator**
    - The tool current remained stable during normal cycles but showed **spikes and higher variability during faults**.
    - Grip loss correlated with inconsistent or unstable tool current, while protective stops often coincided with sharp current surges.

- Rolling averages and fault overlays confirmed that **higher tool current correlates with increased fault probability**.

5. **Outlier & Anomaly Patterns**

   - Fault cycles contained **significantly more outliers** (often 200+) compared to normal operation (~80–100).

   - **Joint speeds** contributed the most outliers, especially at J2, J3, and J5, reinforcing the role of motion irregularities in triggering faults.

6. **Fault Event Characteristics**

   - **Protective Stops** were short-lived (5–11s), not strongly linked to overheating, but triggered by irregular load.

   - **Grip Losses** were also short (9–18s), with evidence that the robot applied more current to maintain grip but failed.

   - In contrast, **normal operations ran for several minutes stably**, showing clear separation between fault and non-fault behavior.

Statistical Analysis, Threshold Determination, and Data Overlap

1. Splitting by Joint and Target

To isolate the behavior of each sensor, the dataset was split per joint (J0–J5) and per fault type (Protective Stop, Grip Loss). This allowed us to analyze each signal's contribution to faults independently rather than across the entire robot.

Note: All the analysis is coming from a joint 0 and specifically with the protective stop

2. Statistical Thresholds (Youden's Index & Quartile Analysis)

We applied **Youden's Index (J = Sensitivity + Specificity − 1)** on the ROC curve to identify **optimal fault thresholds** for each feature.

Example thresholds from Joint 0 (Protective Stop):

- **Temperature_T0 > 34.2°C** → Fault risk increases
- **Current_J0 > −0.067** → Higher fault likelihood
- **Speed_J0 < −0.004** → Possible fault indicator

Additionally, **quartile-based fault rates** were calculated to observe how fault probability rises across feature ranges.

Conclusion:

- Fault risk generally increases in the upper quartiles of temperature.
- Speed faults are concentrated in the lower quartiles (negative speeds).

3. Data Overlap and Density Analysis

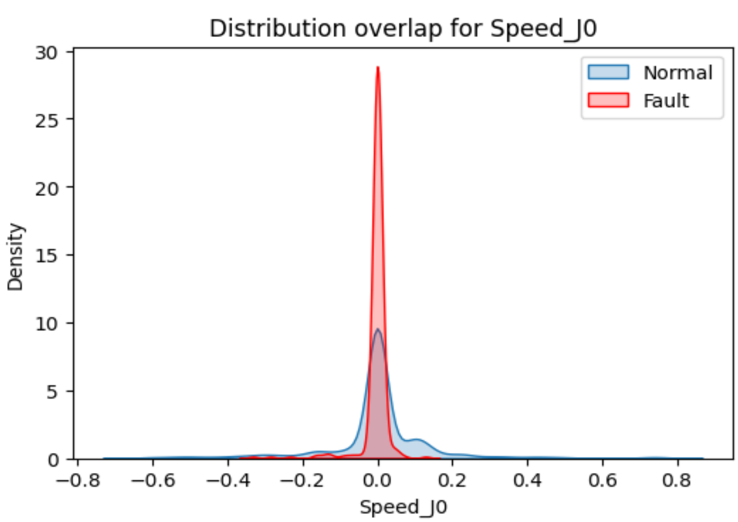
To visualize **normal vs fault distributions**, we used:

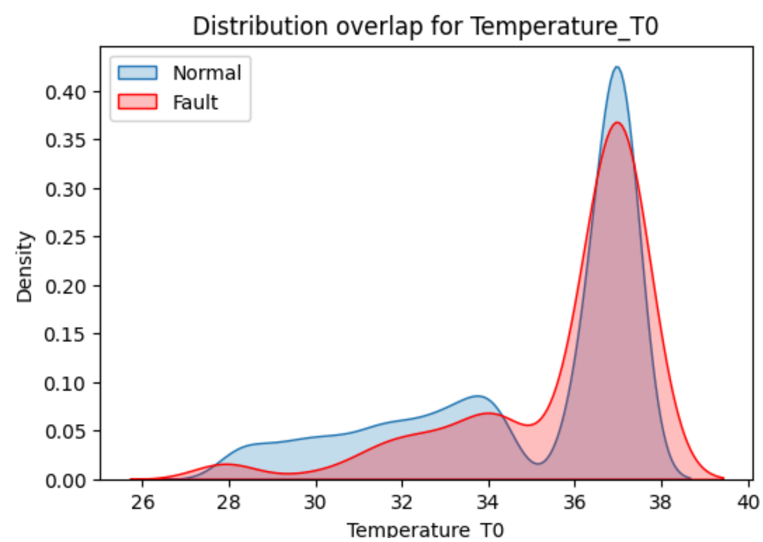
- **Kernel Density Estimates (KDEs)** for each feature
- **Kolmogorov-Smirnov (KS) tests** for distributional differences

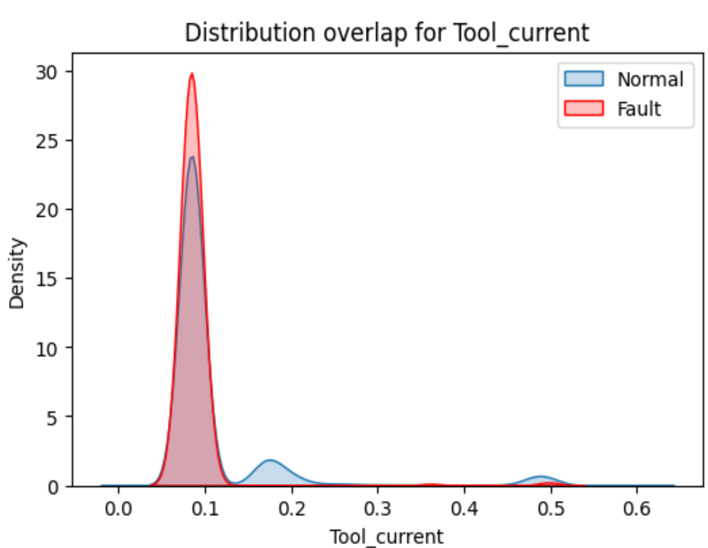
Findings:

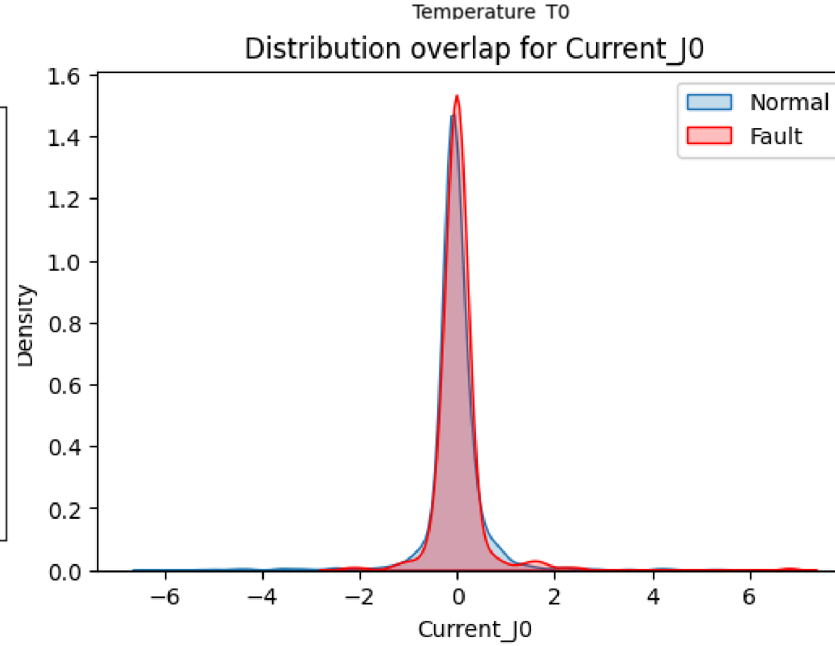
- Fault and normal data distributions **heavily overlap** for all features.
- KS statistics were low, confirming **no clear separation** between fault and normal classes.

As a result, **simple single-feature thresholds** inevitably misclassify many normal points as faults, especially with highly imbalanced data.









4. Effect of Data Imbalance

- Faults constitute **~3–4% of all data**.
- Even small overlaps create many **false positives**, lowering precision while recall remains high.

Example:

- High recall: Most faults detected.
- Low precision: Many normal points are misclassified as faults.

5. Threshold Validation with Classification Metrics

To test threshold performance, we computed:

- **Confusion matrices**
- **Precision, Recall, and F1-scores**

Results showed:

- **High recall but low precision** thresholds detect most faults but generate many false alarms.
- Confirms overlapping data distributions limit single-threshold accuracy.

6. Key Insights

- **Temperature and speed** have the most significant influence on fault occurrence. However, thresholds alone lack predictive precision. At the plot with the density of all features, faults appear most frequently at high temperatures of 36 °C and high vibration speeds.
- **Overlap + imbalance** explains low precision even when recall is acceptable.