

Chatter Detection in Simulated Machining Data: A Simple Refined Approach to Vibration Data

Matthew Alberts^{1*†}, Jamie Coble^{2*†}, Bradley Jared^{3†}, Jaydeep Karandikar^{4†},
Anahita Khojandi^{1†}, Tony Schmitz^{4†}, Sam St. John^{1†}

^{1*}Department of Industrial and Systems Engineering, University of Tennessee, 851
Neyland Drive, Knoxville, 37996, TN, USA.

^{2*}Department of Nuclear Engineering, University of Tennessee, 863 Neyland Drive,
Knoxville, 37996, TN, USA.

³Department of Mechanical, Aerospace, and Biomedical Engineering, University of
Tennessee, 1512 Middle Drive, Knoxville, 37996, TN, USA.

⁴Manufacturing Science Division, Oak Ridge National Laboratory, 1 Bethel Valley Road,
Oak Ridge, 37830, TN, USA.

*Corresponding author(s). E-mail(s): malberts@vols.utk.edu; jcoble1@utk.edu;
Contributing authors: bhjared@utk.edu; karandikarjm@ornl.gov; khojandi@utk.edu;
tony.schmitz@utk.edu; [sstjohn3@vols.utk.edu](mailto:ssjohn3@vols.utk.edu);

[†]These authors contributed equally to this work.

Abstract

Vibration monitoring is a critical aspect of assessing the health and performance of machinery and industrial processes. This study explores the application of machine learning techniques, specifically the Random Forest (RF) classification model, to predict and classify chatter—a detrimental self-excited vibration phenomenon—during machining operations. While sophisticated methods have been employed to address chatter, this research investigates the efficacy of a novel approach to an RF model. The study leverages simulated vibration data, bypassing resource-intensive real-world data collection, to develop a versatile chatter detection model applicable across diverse machining configurations. The feature extraction process combines time-series features and Fast Fourier Transform (FFT) data features, streamlining the model while addressing challenges posed by feature selection. By focusing on the RF model’s simplicity and efficiency, this research advances chatter detection techniques, offering a practical tool with improved generalizability, computational efficiency, and ease of interpretation. The study demonstrates that innovation can reside in simplicity, opening avenues for wider applicability and accelerated progress in the machining industry.

Keywords: random forest, machine learning, chatter, stability, recursive feature elimination simulation, additive manufacturing

This manuscript has been authored in part by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the US

1 Introduction

Vibration monitoring has been an integral practice in assessing the health of rotating machinery and industrial processes for decades. Among various applications of vibration-based damage detection, monitoring rotating machinery has stood out as one of the most mature and successful endeavors [1]. While the field of anomaly detection boasts an array of complex methodologies and advanced algorithms, it is crucial not to overlook the simplicity and efficiency embodied in the Random Forest (RF) model.

In recent years, the machining industry has increasingly turned to vibration monitoring to evaluate machine health and optimize processes [2]. The vibrations emitted by machinery serve as early indicators, empowering operators to make critical decisions during machining operations or even before their commencement [3]. Yet, the challenge of chatter, a disruptive self-excited vibration occurring between the tool and the workpiece, has persisted in machining for over a century [4]. Chatter results in suboptimal surface finishes, reduced material removal rates, and potential damage to both machine tools and workpieces [5].

Our research offers a fresh perspective on this age-old problem: the power of a meticulously designed RF classification model, emphasizing simplicity, as an effective alternative. This model can be easily comprehended and swiftly deployed on the shop floor.

This study makes a distinct contribution to the ongoing research in machining by exploring the potential of vibration-based features combined with time-series features as classifiers for the identification and prediction of chatter in machining processes. In a landscape where intricate methods dominate, our research boldly champions the idea of harnessing the straightforward RF classification model. Building on the foundation laid by Shevchik et al. in their predictive application

of random forest models [6], this study acknowledges the existence of more advanced methodologies. However, our emphasis lies in demonstrating that the dependable RF model can yield remarkable results, especially in industrial settings where accessibility, speed, and interoperability are paramount. In an era that prizes computational speed and practicality, our research sets out to prove that innovation does not always necessitate complexity.

This perspective aligns chatter, a fault detectable through vibration data, with conventional fault diagnosis techniques, which typically entail feature extraction and fault identification [7]. Although a range of signal processing techniques, including statistical analysis [8], Fast Fourier Transforms (FFT) [9], wavelet packet transform (WT) [10], empirical mode decomposition (EMD) [11], and sparse representation methods [12], have been proposed for feature extraction, these methods often require substantial prior knowledge and may have limitations in complex, dynamic working environments.

This research seeks to bridge the gap between advanced theory and practical application. By demonstrating that a simplified RF model can effectively classify chatter and aspire to make this valuable tool more accessible to machinists and industry professionals. This approach retains the merits of precision and reliability while streamlining the process, reducing the reliance on extensive computational resources and specialized expertise.

The subsequent sections will delve into the methodology, experimental results, and implications of our findings. While the chatter detection field has seen the advent of more complex techniques, this research will show that simplicity, as exemplified in the "keep it simple" approach, can be a form of sophistication that catalyzes broader applicability and accelerates progress in the machining industry.

1.1 Feature Extraction Approaches

Despite the plethora of feature extraction methods available for chatter prediction in machining, our research introduces a nuanced methodology that elevates the efficacy of these techniques. Traditional feature extraction methods have laid a strong foundation, yet they often fall short when

Department of Energy (DOE). The US government retains and the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for US government purposes. DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (<http://energy.gov/downloads/doe-public-access-plan>).

addressing the dynamic and non-linear characteristics of chatter. In contrast, our approach synergistically integrates time-series analysis with Random Forest (RF) classifiers, presenting an innovative solution that simplifies the complex landscape of anomaly detection. This integration not only enhances the precision of chatter detection but also paves the way for a model that is intuitively interpretable and rapidly deployable in industrial settings.

1.1.1 Traditional Feature Extraction Methods

Traditional feature extraction methods such as Autocorrelation (AC), Power Spectral Density (PSD), and Fast Fourier Transform (FFT) have been the cornerstone of vibration data analysis. Despite their widespread application, these methods encounter challenges in handling the intricate nature of machining vibrations. Our study builds upon these conventional techniques by refining the feature selection process, thus enabling the detection of chatter with heightened sensitivity and specificity. Previous researchers have also explored combining multiple traditional feature extraction methods. For instance, Yesilli et al. utilized coordinated peaks in AC, PSD, and FFT plots as features for chatter classification in turning [13]. However, determining which peaks are meaningful in FFT can be challenging due to signal variability across different machine setups and machining parameters. This study addresses the limitations inherent in traditional methods, such as their inability to process non-stationary signals, by implementing a tailored RF model adept at navigating the complexities of machining data.

1.1.2 Advanced Feature Extraction Techniques

In recent years, a plethora of feature extraction methods for the prediction of chatter in boring processes has been introduced. Techniques such as Empirical Mode Decomposition (EMD) and Support Vector Machine (SVM) classifiers have shown promise. Li et al. proposed an EMD-based approach, breaking down vibration signals into intrinsic mode functions (IMFs) and creating feature vectors for these IMFs [14]. Similarly, Chen and Zheng utilized SVM in tandem with Recursive Feature Extraction to detect chatter in

milling operations [15]. Despite their effectiveness, these methods often require extensive manual pre-processing and suffer from computational inefficiencies, leading to potential drawbacks such as the need for identifying informative decompositions manually [16][17].

Moreover, additional research has embraced sophisticated techniques like stacked denoising autoencoders, entropy, and coarse-grained information for feature extraction in various machining and grinding scenarios [18][19]. Notably, Li et al. combined multiscale power spectral entropy (MPSE), multiscale permutation entropy (MPE), and Laplacian scores to forecast chatter in milling operations [20]. While these advanced methods are promising, they are not without challenges, including intricate parameter selection and high computational demands.

To address these challenges, our method introduces a streamlined Random Forest (RF) model that integrates the strengths of advanced feature extraction techniques, such as EMD and SVM, while bypassing their complexities and extensive pre-processing requirements. This innovative approach significantly simplifies the process, eliminating the need for manual intervention and making it a more viable and practical tool for real-time machining diagnostics. Our proposed model stands out for its elegant simplicity, offering a more efficient solution for on-the-floor machining diagnostics.

1.1.3 Specific Application Cases

In the domain of monitoring rotating assets for operational anomalies, the detection of issues such as chatter is paramount. Aslan and Altinas notably proposed the utilization of FFT and spectral data to facilitate online chatter detection from spindle drive motor data [21]. This approach, while innovative, is not without its limitations. Specifically, FFT's inherent assumptions regarding the stationarity and linearity of signals pose significant challenges, as real-world vibration signals from such machinery often exhibit non-stationary and transient characteristics, potentially compromising the accuracy of FFT-based methods. Another method that is focused on the spindle is that of Spindle Speed Variation (SSV). SSV is meant to vary the speed of the spindle should chatter be detected, this does require a

machine with a SSV setup and is not found on most milling machines in additive manufacturing environments. Li et al, also mentioned the possible adverse effects of speed variation on tool wear, power fluctuation, and harmful overheating [22]. This infers that modeling milling data and predicting chatter before machining begins would be optimal.

Recognizing these limitations, our research pivots towards adapting the Random Forest (RF) model to better align with the dynamic nature of machine vibrations. By accounting for the inherently transient and non-stationary characteristics of these signals, our adapted RF model offers a more accurate and reliable mechanism for real-time chatter detection, particularly in critical applications like spindle drive monitoring. This enhanced adaptability not only improves the detection accuracy but also ensures seamless integration with existing monitoring systems. By doing so, our model demonstrates the considerable potential to transform chatter detection across various machining operations, setting a new standard for precision and reliability in the field.

1.1.4 Challenges and Limitations

Feature extraction methods are fundamental in the analysis of complex data sets, yet each method inherently carries its own set of distinct challenges. These challenges range from the intricacies of parameter selection and the sensitivity to noise in the data, to the interpretation of results and the scalability issues that emerge when dealing with large datasets harboring numerous features. Such complexities often hinder the effectiveness and applicability of these methods in real-world scenarios, particularly in the vibration analysis of milling operations.

Recognizing these challenges, our research delves into a systematic approach aimed at mitigating these obstacles. The subsequent sections will introduce and detail our proposed methodology, which seeks to capitalize on the inherent strengths of the Random Forest (RF) model. This model is renowned for its robustness against noise and its proficient ability to evaluate the importance of features that are instrumental in enhancing the generalizability of the analysis [23]. By harnessing these capabilities, our method aims

to transcend the traditional limitations associated with feature extraction techniques.

Our approach does not merely stop at identifying the challenges; it proactively addresses them by integrating the simplicity and efficiency of the RF model with advanced analytical strategies. This integration results in a sophisticated yet practical tool, one that is adept at managing the nuances of parameter selection, reducing the impact of noise on the data, and interpreting complex results in a meaningful way. Furthermore, our method significantly improves scalability, enabling the analysis of extensive datasets without compromising on accuracy or computational efficiency. Through these enhancements, our proposed approach promises to revolutionize the vibration analysis of milling operations, offering a model that is not only grounded in solid theoretical foundations but also distinguished by its applicability and effectiveness in real-world industrial settings.

1.2 Approach and Research Contributions

This comprehensive study introduces a streamlined and innovative approach for the prediction and detection of chatter in milling operations. Central to our methodology is the integration of simulated data, advanced time-series feature extraction using Fast Fourier Transform (FFT) data features, and the robust classification capabilities of a Random Forest (RF) model. This multifaceted approach is designed not only to enhance the precision of chatter detection but also to address and overcome the common challenges associated with limited real-world datasets.

Our primary contribution is the strategic utilization of simulated vibration data, which serves as a powerful tool to circumvent the constraints of small-scale real-world datasets. Feng et al. showcased how similar grinding topography simulated data could be to real-world data supporting this study's use of simulated matching data to be an actual representation of real-world data [24]. This innovative use of simulated data significantly expands the versatility and applicability of our chatter detection model, making it suitable for a wide array of machining scenarios. By employing

FFT for vibration data feature extraction, a well-established method that is recognized and valued in the industry for its reliability and efficacy.

Furthermore, our RF classification model stands out for its unique ability to harness the potential of this simulated data, leading to a solution that not only reduces the dependency on extensive data collection from physical experiments but also exhibits exceptional adaptability to various machining conditions. The incorporation of FFT and time-series features into our model ensures computational efficiency and produces results that are readily interpretable by industry professionals, thus bridging the gap between theoretical research and practical application.

The experimental validation of our approach, which is detailed in the subsequent sections of this study, provides a clear demonstration of our model’s capability to accurately emulate real-world machining scenarios. This validation firmly establishes the practicality and applicability of our model in industrial settings, offering a promising solution to the pervasive challenge of chatter prediction in milling operations. By marrying the strengths of simulated data, FFT feature extraction, and the RF classification model, this study presents a holistic approach that is not only innovative in its design but also exemplary in its performance and relevance to the industry.

1.2.1 Leveraging Simulated Vibration Data

Our approach distinguishes itself by harnessing the power of simulated vibration data. This novel strategy obviates the need for resource-intensive and time-consuming real-world data collection, a common hurdle in previous studies. Moreover, our model’s adaptability extends beyond specific machining configurations, offering the capability to extrapolate to diverse setups. This work builds upon the foundation laid by prior research, such as the investigations conducted by Yesilli and Khasawneh, which employed traditional feature extraction techniques like FFT, autocorrelation, and power spectral density [25].

1.2.2 Streamlined Feature Selection for Enhanced Model Efficiency

While earlier studies employed a range of features, including FFT, MPSE, MPE, or Laplacian scores,

this study streamlined the approach by focusing solely on FFT and time-series data features. This strategic simplification aims to address the challenges associated with feature selection in chatter classification models. By reducing the feature set, the model’s comprehensibility is enhanced, computational demands are reduced, and the risk of overfitting the training data is mitigated. This aligns with Dutta et al. who saw RF have a more than 30% reduction in computation time and power to other methods [26]. This reduction aligns with our goal of achieving model generalizability, a principle emphasized by Jia et al. [27].

1.2.3 Experimental Validation

This section outlines the experimental setup and the validation process for assessing the efficacy of the RF model in classifying chatter in machining processes. The simulated data generation process will be described and how this synthetic dataset emulates real-world machining conditions will be explained. Also, the details of the metrics and criteria that will be employed to evaluate the performance of our RF model, including accuracy, precision, recall, and F1 score.

In summary, this study presents an efficient and versatile approach to chatter prediction, empowered by simulated data, FFT feature extraction, and the RF classification model. By emphasizing simplicity and feature selection precision, this study not only advances the field of chatter detection but also offers a practical tool with broad applicability, simplified interpretability, and enhanced generalizability, thus addressing the current limitations and complexities inherent in the domain.

2 Materials and Methods

This investigation unfolds an innovative methodology designed to predict chatter in milling operations with a pioneering integration of simulated vibration data and sophisticated time-series feature extraction techniques, with a particular emphasis on FFT data features. Our methodological innovation lies in the utilization of a Random Forest (RF) classification model, conceived to circumvent the constraints typically posed by limited real-world datasets. Such constraints often hinder

the scalability and flexibility required for a comprehensive chatter detection model. This study demonstrates the unique potential of simulated datasets to replicate a diverse array of milling scenarios, thus broadening the applicability of the RF model in practical settings.

2.1 Cross-Validation Techniques and Dataset Size Importance

To bolster the robustness and reliability of our predictive model, our research meticulously adopts k-fold cross-validation techniques, a method celebrated for its effectiveness in evaluating model performance. This approach provides a rigorous framework for assessing the model’s robustness, significantly mitigating risks of overfitting by partitioning the data into training and testing subsets multiple times. Such a meticulous evaluation allows for a thorough understanding of the model’s predictive capabilities, ensuring that its performance is not just a result of chance or overfitting to a particular set of data.

Furthermore, this study recognizes and emphasizes the pivotal role of dataset size and diversity in the performance of machine learning models. Large and diverse datasets are instrumental in enhancing the model’s ability to generalize and make accurate predictions across various scenarios. In this study, the substantial and carefully curated dataset, especially with the deliberate exclusion of extrapolated data from the training set, plays a crucial role. This strategic dataset composition reinforces the model’s proficiency, enabling it to operate effectively and consistently across a broad spectrum of machining conditions. Through this combination of rigorous validation techniques and strategic dataset management, our approach ensures the utmost robustness and reliability of the predictive model in practical, real-world applications.

2.2 Data Description and Dataset Division

This comprehensive study delves into the intricacies of machining dynamics by employing a state-of-the-art simulated vibration data source. This data is meticulously generated from a sophisticated machining simulation model, a brainchild of Schmitz and Smith as detailed in their seminal

work, "Machining Dynamics: Frequency Response to Improved Productivity" [5]. This model is not just a mere representation but a close emulation of the conditions experienced by a single milling machine, as visually encapsulated in Figure 1.

The depth of this simulation is further exemplified through Figures A1 and A2, which provide a stark visual contrast between the raw vibration acceleration data for datasets characterized by the presence and absence of chatter. A notable observation from this comparative analysis is the marked tenfold decrease in acceleration values during instances of chatter, a testament to the precision and authenticity of the simulation.

Moreover, the subtle intricacies of this simulation are brought to the forefront in Figures A3 and A4. These figures meticulously illustrate data runs that teeter on the delicate threshold between stability and instability, a domain notoriously challenging for chatter prediction. This can best be seen in Figure 2, where the instability zone is between the chatter and no chatter areas. It is within this complex landscape that the prowess of Fast Fourier Transform (FFT) and time-series features, particularly those extracted via TSFresh, becomes indispensable. The FFT, with its ability to spotlight peak values at specific frequencies, emerges as a critical tool in identifying potential machine issues. When synergized with the nuanced insights provided by time-series features, a comprehensive framework for effective chatter classification is established, extending its applicability not only within but also beyond the confines of the training datasets A and A0. Doing this will allow the model to include the use of nonlinear parameters to evaluate the degree of randomness, complexity or similarity of the signals [28].

This simulation, with its detailed visual representations and the integration of FFT and time-series features, underscores the nuanced differences in vibration acceleration between stable operations and those prone to chatter. The precision and depth of this simulated data echo the operational realities of a milling machine under varied conditions, offering an unprecedented level of detail and accuracy. Consequently, this study not only leverages but also significantly enriches the foundational work of Schmitz and Smith, offering novel insights and methodologies that hold the potential to transform the field of machining dynamics.

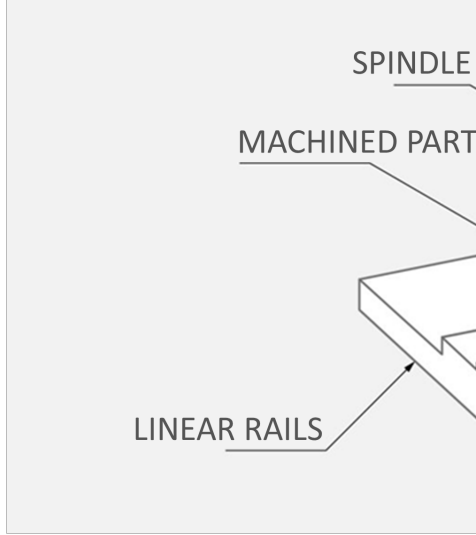


Fig. 1: Example Milling Machine Setup.

2.3 Data Sets A and A0

Data sets A and A0 are derived from the same machining tool setup, with variations in cutting depth and rotational speed of the machine tool head. The cutting parameters remain consistent, with a cutting force of $700 \times 10^6 N/m^2$, feed per tooth of 0.1 mm/tooth, and a tool diameter of 10 mm equipped with four cutting teeth. Set A consists of 4000 data files, covering RPM settings ranging from 6,000 to 10,000 rpm in 50 rpm increments, with the cutting depth varying from 0.2 to 10 mm in 0.2 mm increments. Set A0 comprises 2500 data files, representing RPM settings from 10,050 to 12,000 rpm in 50 rpm increments, with the cutting depth varying from 10.2 to 15 mm in 0.2 mm increments. Notably, Set A0 serves as an extrapolated data set and is excluded from the training set for the RF classifier model. In Figure 2, the example dataset will first see chatter at a cutting depth of around 6mm, this is in conjunction with the specific machine setup and will vary based on inputs like cutting teeth, cutting force, material, rpm, machine dampening devices, etc.

2.4 Feature Extraction and Selection

In our goal to refine the predictive capabilities of our model while ensuring its operational efficiency, the method of Recursive Feature Elimination (RFE) has been strategically incorporated into this model. This technique stands at the forefront

of our feature selection process, iteratively pruning the less significant features from our initial comprehensive set. The essence of RFE lies in its ability to distill the core, most impactful features, thereby optimizing the model's interpretability and significantly reducing computational demands [20][30]. This rigorous approach not only sharpens the model's focus on the most influential attributes but also serves as a preventive measure against the risk of overfitting, ensuring robust and reliable predictions.

The application of RFE in our methodology has been instrumental in achieving a streamlined feature set, intricately documented in our detailed tables. Table 1 is a testament to BEDs, presenting the FFT features meticulously extracted from each dataset, capturing the frequency-based nuances integral to our analysis. Complementing this, Table 2 offers an exhaustive overview of the time-series features extracted from each dataset, illustrating the rich temporal patterns and dynamics encapsulated in our data. These tables not only serve as a repository of the refined features but also underscore the thoroughness of our feature selection process.

The result of this feature selection process is a refined and optimized set of FFT and time-series features, which form the foundational inputs for our Random Forest (RF) classification model. By harnessing RFE, it ensures that the RF model is fed with only the most crucial features, significantly enhancing the model's interpretability and computational efficiency. This strategic distillation of features sets the stage for a streamlined and effective RF classification model, poised to deliver precise and actionable insights. In essence, the integration of RFE in our feature selection process reflects our unwavering commitment to precision, efficiency, and clarity, driving the development of a robust and interpretable predictive model for our analytical endeavors.

2.5 Data Preprocessing

In the data preprocessing phase of this study, a structured approach was adopted for feature extraction from the collected raw vibration data during machining operations. Custom Python scripts were meticulously developed for the extraction of Fast Fourier Transform (FFT) features. Concurrently, the TSFresh library was utilized for

Table 1: Set A0 and A0 FFT Features Extracted.

Feature	Description
Acceleration Peak (g)	$\text{peak} = \max(a)$
Acceleration RMS (g)	$RMS = \sqrt{(1/n) \sum_{i=1}^n a_i^2}$
Crest Factor	$Crest = X_{peak} / X_{RMS}$
Standard Deviation (g)	$\text{std} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}$
Velocity RMS (in/s)	$V_{RMS} = \sqrt{A^2/2}$
Displacement RMS (in)	$D_{RMS} = \sqrt{(a/(2 * \pi * f)^2)^2}$
Peak Frequency (Hz)	Highest Peak Frequency (Hz)
RMS (g) from 1 to 65hz	$RMS_{1-65} = \sqrt{(1/n) \sum_{i=1}^{65} a_i^2}$
RMS (g) from 65 to 300hz	$RMS_{65-300} = \sqrt{(1/n) \sum_{i=65}^{300} a_i^2}$
RMS (g) from 300 to 6000hz	$RMS_{300-1000} = \sqrt{(1/n) \sum_{i=300}^{1000} a_i^2}$

the extraction of time-series features, reflecting a systematic integration of computational techniques [31]. The transformation of raw data sets A and A0 into feature data sets B and B0 was methodically executed through the application of TSfresh and FFT methodologies. The challenges associated with discerning stability and instability from FFT data alone are highlighted in Figures A5, A6, A7, and A8.

Table 2 details the time-series features extracted from each data set, culminating in the generation of 846 features. Recognizing the potential risk of model overfitting and loss of generalizability due to the extensive number of features, a process of Recursive Feature Elimination (RFE) was employed. This process effectively distilled the feature set to 17 most significant features for the RF model, enhancing the model’s predictive accuracy while maintaining its computational efficiency. These 17 features were categorized into two distinct groups: 10 FFT features and 7 time-series features, each set contributing uniquely to the model’s performance. The refined feature data sets are designated to be utilized as input for the RF classification model.

The transformation of raw datasets into precisely refined feature sets was achieved through the diligent application of custom Python scripts

alongside the TSfresh library, ensuring a meticulous analysis of each feature’s significance in the context of machine vibration analysis. The preprocessing phase is critically underscored as instrumental in accurately identifying the markers of stability and instability within the datasets, as depicted in our figures. This systematic approach to feature extraction and selection forms the cornerstone of the RF classification model, ensuring a robust analytical framework for the prediction of chatter in milling operations.

2.6 Model Development

In the initial phase of feature engineering, a comprehensive correlation analysis was undertaken to discern the predictive value of the generated features in the context of the vibration data acquired during machining processes. Notable among these features were the aggregated Fast Fourier Transform (FFT) outputs, along with statistical measures such as the spectral centroid, mean, variance, skewness, and kurtosis of the absolute Fourier spectrum. Despite preliminary indications suggesting their potential utility, further analysis relegated these features to the non-predictive category.

Table 2: Set A0 and A0 TSFresh Features Extracted.

Feature	Description
Ratio value number to time series length	The number of unique values versus the total number of values
Benford correlation	How often a value starts with a certain number, in analytics overwhelmingly a value is most likely to start with a 1
Change quant f-aggr "var" False qh 1.0 ql 0.4	Aggregator function of the differences taken over a specific range of upper and lower quartiles
FFT coefficient attr "imag" coeff 55	Fast Fourier Transform of the imaginary part of the data with a coefficient of 55
FFT coefficient attr "imag" coeff 77	Fast Fourier Transform of the imaginary part of the data with a coefficient of 77
Agg linear trend "stderr" len 10 f agg "min"	Linear least squares regression for certain attributes for a certain number of time series data points
Permutation entropy dimension 4 tau 1	Counts the frequency of permutation and returns the appropriate entropy, this is a complexity measure for time series data

Proceeding to the feature elimination phase, the Recursive Feature Elimination (RFE) strategy was employed with methodical precision. This strategy facilitated the exclusion of non-contributory features, thereby refining the feature set to a succinct ensemble that is computationally less demanding and more aligned with the predictive exigencies of the model.

The development of the model, as systematically outlined in Figure A9, underscores our exacting approach to feature selection. This process was instrumental in ensuring the retention of only those features that exhibited substantial predictive merit. The RF analysis was characterized by judicious employment of various training-to-test data splits, encompassing ratios of 70-30, 50-50, and 40-60, to validate the model's predictive robustness across a spectrum of scenarios.

In essence, this study's methodology exemplifies a paradigm of structured feature selection that is critical to the construction of an accurate and computationally efficient RF classification model. The approach delineates a clear trajectory from the identification through to the elimination of non-essential features, thus sharpening the model's predictive acuity and enhancing its operational expediency. This structured approach not only highlights the paramount importance of feature relevance in RF analysis but also exemplifies the adaptability and analytical rigor requisite for handling complex machining data.

2.7 Evaluation Metrics

The framework of this study delineates the suite of evaluation metrics that have been systematically employed to assess the performance of our

Random Forest (RF) model in the prediction of chatter within milling operations. The explication of these metrics is paramount for a comprehensive understanding of the model’s effectiveness and its connection to practical implications within the operational environment.

Central to our evaluative strategy is the employment of the Area Under the Curve (AUC) metric. The AUC metric stands as a pivotal and robust indicator of the model’s classification accuracy, offering an integrative assessment that considers the entirety of possible classification thresholds. AUC values range from 0.5 to 1.0. An AUC of 0.5 suggests that the model performs no better than random chance, while an AUC of 1.0 indicates perfect classification. Values greater than 0.5 demonstrate the model’s ability to distinguish chatter from no chatter with better-than-random accuracy. Specifically, it quantifies the percentage at which the model correctly predicts chatter occurrences. For instance, a model with a 97% AUC score signifies a 97% accuracy rate in predicting chatter. This assessment allows us to ascertain the model’s generality and its potential to be effectively deployed across a broader spectrum of machining speeds and cutting depths. The simulated data contains a chatter metric that indicates whether or not the dataset experienced chatter. This can be seen in Figure 2 below. From this, you will see the line between chatter and no chatter. It will be in this zone that will be the most difficult to predict chatter or no chatter, given the minute changes in variables that can influence whether or not chatter is experienced.

In concert with the AUC, the robustness of the model’s predictive validity is further reinforced through the implementation of k-fold cross-validation. This methodological approach enables a comprehensive validation of the model’s performance by systematically partitioning the data into subsets, thus ensuring that the evaluation is reflective of the model’s ability to generalize across various subsets of data.

Moreover, the significance of dataset size in the context of model evaluation is highlighted, recognizing that the volume and diversity of data are critical to the model’s learning and subsequent predictive performance. By incorporating a substantial and diverse dataset, it ensures that the model’s utility is not confined to theoretical

or controlled scenarios but extends to practical, real-world machining environments.

In summary, the evaluation framework adopted herein is both rigorous and multifaceted, designed to thoroughly scrutinize the model’s efficacy. It is through this comprehensive battery of metrics, with the AUC as the cornerstone of our evaluation, supported by k-fold cross-validation and a substantive dataset, that this study affirms the robustness and practical utility of our RF model in the realm of chatter prediction for milling operations.

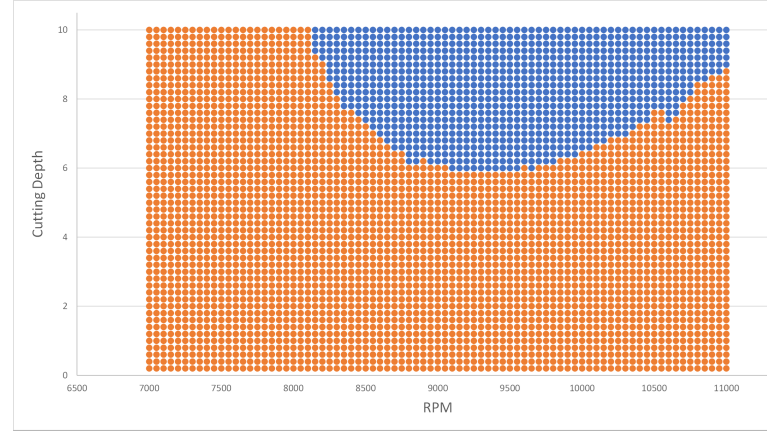


Fig. 2: Simulated Data Curve.

3 Investigation

This study undertakes a methodological investigation through a series of three distinct analyses designed to rigorously assess the robustness of key vibration and time-series features, evaluate the predictive acumen of RF models within diverse datasets, and scrutinize the generalizability of the models with respect to extrapolated data.

Study 1: Consistency of Key Features within Set $i \in \{A, A0\}$

Objective: This analytical segment is devoted to clarifying the consistency of salient features in machining vibration data across identical tool configurations.

1. Feature Identification and Evaluation: The study commences with the identification and subsequent evaluation of paramount features,

denoted by the set $t F_i$, , within each dataset $i \in \{A, A0\}$

2. Intersection of Feature Sets: The investigation proceeds to discern the intersection of feature sets, represented by $F_A \cap F_{A0}$, to gauge the uniformity of feature significance across datasets.
3. Quantification of Common Features: The cardinality, denoted as $n(F_A \cap F_{A0})$, quantifies the common features, thereby providing a metric for consistency across the datasets.

Study 2: Generalizability of RF Models within Set $i \in \{A, A0\}$

Objective: This phase aims to assess the generalizability of the RF models in predicting chatter across identical tool configurations, thereby establishing the models' predictive stability.

1. RF Model Construction: Individual RF models are meticulously constructed for each dataset $i \in \{A, A0\}$
2. Intra-dataset Performance Assessment: The predictive performance of each model within its respective dataset is rigorously evaluated using the Area Under the Curve (AUC) metric, denoted as $AUC_{A,A}^c$ and $AUC_{A0,A0}^c$.
3. Cross-dataset Performance Evaluation: The models' predictive acumen is further scrutinized across datasets $i \in A, A0$ employing the AUC metric, represented as $AUC_{cA,A0}$ and $AUC_{A0,A}$, ensuring a comprehensive assessment of the models' robustness.

Study 3: Generalizability of RF Models in Extrapolated Data

Objective: This segment explores the capability of RF models to sustain predictive accuracy and reliability when applied to extrapolated data, thereby testing the models' adaptability in broader application scenarios.

1. Model Development for Extrapolated Data: RF models are adeptly developed for sets $i \in \{A, A0\}$ and $j \in \{B, B0\}$, catering to both original and extrapolated data contexts.
2. Performance Appraisal for Extrapolated Data: The models' performance within each dataset is scrupulously measured using the AUC metric, represented as $AUC_{B,B}$, for dataset $j \in \{B, B0\}$.
3. Comprehensive Evaluation of Generalizability: An extensive assessment of the models' performance is conducted across all datasets

$i \in A, A0$ and sets $j \in B, B0$ utilizing the AUC metric, denoted as $AUC_{A,B}$, $AUC_{A0,B}$, $AUC_{B,A}$, and $AUC_{cB,A0}$, thus affirming the models' adaptability and forecasting potential.

This systematic and rigorous investigation ensures an in-depth examination of feature consistency, model predictiveness, and generalizability, contributing invaluable insights into the potential of RF models to advance chatter prediction in machining operations.

4 Results

This section delineates the findings of our comprehensive investigation, each study contributing nuanced insights into the effectiveness, precision, and adaptability of our RF model in the context of chatter prediction.

4.1 Features and Descriptive Statistics

The initial phase of feature extraction, facilitated by the TSFresh Python package, yielded an exhaustive array of 846 features, spanning both time-series and FFT-based vibration attributes. A methodical feature reduction process was undertaken to distill this extensive set down to 17 paramount features, thereby ensuring the model's efficiency and predictive strength.

4.2 Investigation Results

The investigation unfolded through a series of structured studies, each clarifying distinct facets of the model's performance and the viability of the implemented features.

4.2.1 Study 1 Results

The first analytical study illuminated the internal consistency of key features across sets A and $A0$, corroborating the model's acuity in identifying features indicative of chatter. Figures 3 and 4 elegantly illustrate this congruence, presenting the top features in both vibration and time-series categories for the respective sets.

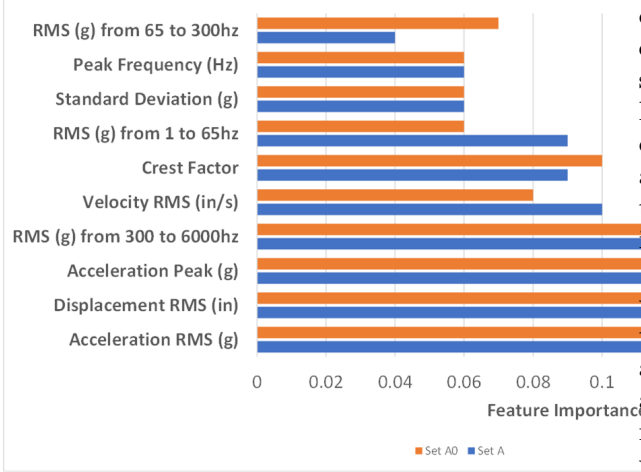


Fig. 3: Set A and A0 ten vibration features.

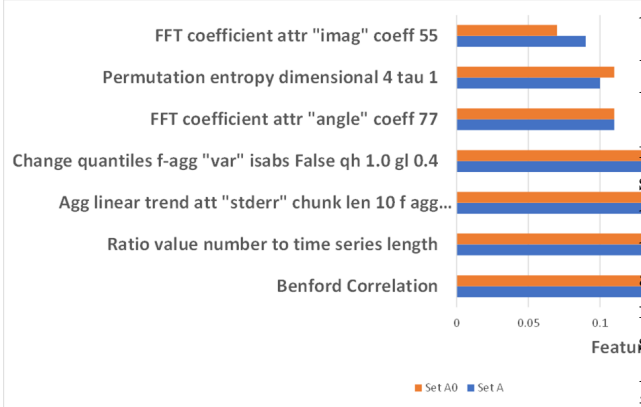


Fig. 4: Set A and A0 time-series features.

4.2.2 Study 2 Results

The analyses conducted in Study 2 furnish compelling evidence regarding the high predictability and generalizability of the RF models across the datasets in classifying chatter. The empirical results for Set A were remarkably robust, with all Area Under the Curve (AUC) values surpassing the threshold of 0.997, indicative of the model's profound predictive precision, as detailed in Table 3. Table 4 details the results for set A0. This level of accuracy establishes the model as exceptionally reliable within the dataset population under study.

The AUC metrics for Set A serve as a testament to the predictability of different feature groups within the datasets. Specifically, the predictability of vibration features within Set A,

denoted as $AUC_{A,A}^{vibe}$, reached the ceiling value of 1.000. In contrast, the predictability of time-series features, denoted as $AUC_{A,A}^{time}$, was slightly lower at 0.995. However, the combined feature predictability, represented by $AUC_{A,A}^{top-15}$, achieved a score of 1.000, confirming that vibration features marginally outperform time-series features in terms of importance for this specific dataset.

Similarly, the results for Set A0 align closely with those of Set A, manifesting high classification accuracy across all examined feature groups and reinforcing the model's capability to generalize effectively. The predictability of vibration features within Set A0, as indicated by $AUC_{A0,A0}^{vibe}$, was reported at 0.992, while the predictability of time-series features within the same set, as indicated by $AUC_{A0,A0}^{time}$, attained the optimal value of 1.000. When considering the combined feature predictability for Set A0, as represented by $AUC_{A0,A0}^{top-15}$, the model still exhibits an impressive AUC value of 0.997.

In summary, the RF model's exemplary performance is underlined by the AUC values, which serve as robust indicators of the model's ability to classify chatter with high accuracy. The consistency of these results across different feature sets and dataset variations underscores the model's remarkable generalizability. The predictive precision of the model, as illustrated by near-perfect AUC values, instills a high degree of confidence in its deployment for practical applications within the domain of chatter classification in milling operations.

4.2.3 Study 3 Results

Study 3 reveals favorable predictability of the RF models within dataset B, as outlined in Table 5. The predictability of vibration features within Set B, represented by $AUC_{B,B}^{vibe}$, is 0.824, while the predictability of time-series features within Set B $AUC_{B,B}^{time}$ is 0.869. The combined feature predictability for Set B $AUC_{B,B}^{top-17}$ stands at 0.879.

Study 3 also demonstrates the RF model's ability to extrapolate data beyond the original training dataset's range. This extrapolation includes machine tool rotations per minute and cutting depth values outside the training dataset's range. Notably, the combination of vibration and

time-series features yields the best generalizability results.

5 Discussion

This section provides a comprehensive interpretation of the outcomes of our investigation, underscoring the significance of the methodological simplicity and operational efficiency inherent in our approach. These attributes not only contribute to the robustness of our model but also render it a pragmatic asset within the realm of industrial machining applications.

Our study leverages the capabilities of Recursive Feature Elimination (RFE) within the framework of Random Forest (RF) models, a methodological choice that substantially streamlines the complexity of vibration data analysis. The initial challenge, presented by the daunting quantity of features generated by the TSFresh package, was surmounted by the judicious application of RFE. This tool proved instrumental in distilling the essence of the dataset, selectively sifting through the multitude of features to identify those with paramount significance for chatter prediction.

The integration of selected features, encompassing both time-series data and traditional FFT-based vibration analysis, has substantially augmented the model’s predictive accuracy and generalizability. This strategic amalgamation has not only enhanced the model’s capabilities in stability classification but has also achieved this enhancement without delving into more intricate and computationally intensive methodologies. Such a balance between simplicity and effectiveness is seldom achieved, underscoring the innovative edge of our approach.

The empirical findings of our research outline a pronounced synergy between feature importance and classification accuracy. This synergy was consistently observed across datasets *A* and *A0*, affirming the adeptness of RFE in unearthing the most predictive features. Furthermore, the application of our RF model to extrapolated data in set *B* demonstrated notable generalizability. Achieving an impressive classification accuracy of 87% on data that extends beyond the initial training scope, the model exhibited a modest decline in performance, a testament to its robustness and adaptability across diverse operational conditions.

The practical implications of our findings are profound. The RF model, harnessed with FFT and time-series features, epitomizes efficiency and deployability. Its capacity for swift integration into the shop floor, coupled with its accessibility to professionals across a spectrum of expertise levels, addresses the pressing demands of the contemporary industrial milieu, which places a premium on agility and practical utility.

6 Conclusion

This investigation concludes with a reflective analysis of the strides made through our methodological approach, characterized by its intrinsic simplicity and operational efficacy. The study has demonstrated the competence of a well-crafted RF classification model in accurately categorizing simulated vibration data for the detection of chatter in machining processes. The dataset, albeit comprehensive within its current scope, encompasses less than 10,000 data files. In light of the potential for augmented model robustness and generalizability, future research endeavors will be directed toward expanding the dataset substantially, aiming to surpass the threshold of 100,000 datasets. This expansion is anticipated to substantially refine the model’s predictive precision and its adeptness in navigating the multifaceted landscape of operational conditions prevalent in additive manufacturing settings.

Furthermore, the union of empirical data with simulated data presents a promising avenue for exploration. The reliance on purely simulated data, while instrumental in establishing the foundational robustness of the model, may not fully encapsulate the stochastic nature of real-world operational settings. To surmount this limitation, forthcoming research initiatives will concentrate on integrating empirical data with simulated datasets. This integrative approach seeks to leverage the granularity and unpredictability inherent in real-world data, thereby enhancing the model’s predictive acumen and its adaptability to diverse operational paradigms.

The ultimate ambition of these future explorations is to facilitate a transition from retrospective to real-time predictive modeling. By harnessing the synergistic potential of voluminous, heterogeneous datasets and real-time data integration, the goal is to evolve a model that not only

Table 3: Set A Random Forest model predictability.

Model Features	AUC	Sensitivity	Specificity
Combinds features	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000
Vibration features	0.995 ± 0.005	0.989 ± 0.013	0.972 ± 0.033
Time-series features	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000

Table 4: Set A0 Random Forest model predictability.

Model Features	AUC	Sensitivity	Specificity
Combined features	0.997 ± 0.003	0.994 ± 0.002	0.990 ± 0.008
Vibration features	0.992 ± 0.008	0.972 ± 0.032	0.957 ± 0.048
Time Series features	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000

prognosticates chatter occurrences with heightened precision but also does so in an online, proactive manner. Such a leap towards real-time predictive modeling is poised to revolutionize the domain of chatter detection, heralding new possibilities for predictive maintenance and operational efficiency in the manufacturing sector.

This research substantiates the viability of a streamlined novel yet insightful approach to technological innovation in the realm of industrial machining. It signals the industry to recognize and adopt this paradigm, which underscores that practical, impactful solutions can indeed emerge from simplicity and strategic ingenuity. As research progresses, our endeavor will remain steadfastly focused on broadening the empirical base of our model, enriching its predictive faculties, and ultimately, endowing the industrial machining domain with a tool of unparalleled value and utility.

Declarations

- Author Contributions-Conceptualization, Matthew Alberts. and Tony Schmitz.; methodology, Matthew Alberts and Jamie Coble.; software, Tony Schmitz.; validation, Tony Schmitz, Jamie Coble. and Anahita Khojandi.; formal analysis, Matthew Alberts; resources, Jaydeep Karandikar; data curation, Matthew Alberts; writing—original draft preparation, Matthew Alberts; writing—review and editing, Tony Schmitz and Jamie Coble; supervision, Bradley Jared;

funding acquisition, Y.Y. All authors have read and agreed to the published version of the manuscript.

- Data Availability-Data is unavailable at this time due to privacy restrictions.
- Code Availability-Code is unavailable at this time due to privacy restrictions
- Conflicts of Interest-Conflicts of Interest: No conflict of interest exist for any of the authors related to the conduction or publication of this work. All stated opinions and conclusions are solely those of the authors.

References

- [1] Farrar, C.R. & Doebling, S.W., *Damage detection II: field applications to large*. In: Silva, J.M.M. and Maia, N.M.M. (eds.), *Modal Analysis and Testing*, Nato Science Series, 1999. Dordrecht, Netherlands: Kluwer Academic Publishers
- [2] Goyal, D. & Pabla, B.S., The vibration monitoring methods and signal processing techniques for structural health monitoring: a review. *Archives of Computational Methods in Engineering* 2016 23(4), pp.585-594.
- [3] Carden, E.P. & Fanning, P., Vibration-based condition monitoring: a review. *Structural health monitoring* 2004, 3(4), pp.355-377.

Table 5: Set *B* Random Forest model predictability.

Model Features	AUC	Sensitivity	Specificity
Combined features	0.879 ± 0.121	0.872 ± 0.151	0.869 ± 0.115
Vibration features	0.824 ± 0.076	0.812 ± 0.180	0.808 ± 0.178
Time Series features	0.869 ± 0.131	0.861 ± 0.142	0.852 ± 0.109

- [4] Quintana, G., & Ciurana, J., Chatter in machining processes: A Review. *International Journal of Machine Tools and Manufacture* 2011, 51(5), 363–376.
- [5]
- [6] Shevchik, S. A., Saeidi, F., Meylan, B., & Wasmer, K. (2017). Prediction of Failure in Lubricated Surfaces Using Acoustic Time–Frequency Features and Random Forest Algorithm. *IEEE Transactions on Industrial Informatics*, 13(4), 1541–1553.
- [7] Wu, S.-D., Wu, P.-H., Wu, C.-W., Ding, J.-J., & Wang, C.-C. (2012). Bearing Fault Diagnosis Based on Multiscale Permutation Entropy and Support Vector Machine. *Entropy*, 14(8), 1343–1356.
- [8] Heng, R. B. W., & Nor, M. J. M. (1998). Statistical analysis of sound and vibration signals for monitoring rolling element bearing condition. *Applied Acoustics*, 53(1-3), 211–226.
- [9] Rao, K. R., Kim, D. N., & Hwang, J. J., Fast Fourier transform : algorithms and applications. Springer.
- [10] Huo, Z., Zhang, Y., Francq, P., Shu, L., & Huang, J. (2017). Incipient Fault Diagnosis of Roller Bearing Using Optimized Wavelet Transform Based Multi-Speed Vibration Signatures. *IEEE Access*, 5, 19442–19456.
- [11] Ma, X., Hu, J., & Zhang, L. (2017). EMD-based online Filtering of Process Data. *Control Engineering Practice*, 62, 79–91.
- [12] Jianchao Yang, Wright, J., Huang, T. S., & Yi Ma. (2010). Image Super-Resolution Via Sparse Representation. *IEEE Transactions on Image Processing*, 19(11), 2861–2873.
- [13] Yesilli, M., & Khasawneh, F., *On Transfer Learning of Traditional Frequency and Time Domain Features in Turning*, 2021. Electrical Engineering and Systems Science
- [14] X. Li, Z. H. Yao, & Z. C. Chen, *An effective EMD-based feature extraction method for boring chatter recognition*, *Applied Mechanics and Materials*, vol. 34-35, pp. 1058–1063, Oct 2010.
- [15] G. S. Chen & Q. Z. Zheng, *Online chatter detection of the end milling based on wavelet packet transform and support vector machine recursive feature elimination*, *The International Journal of Advanced Manufacturing Technology*, vol. 95, pp. 775–784, nov 2017.
- [16] Konar, P., & Chattopadhyay, P. (2011). Bearing fault detection of induction motor using wavelet and Support Vector Machines (SVMs). *Applied Soft Computing*, 11(6), 4203–4211.
- [17] Yesilli, M. C., Khasawneh, F. A., & Mann, B. P., *Transfer learning for autonomous chatter detection in machining*. 2022. *Journal of Manufacturing Processes*, 80, 1–27.
- [18] S. Wan, X. Li, Y. Yin, & J. Hong, *Milling chatter detection by multi-feature fusion and adaboostsvm*, *Mechanical Systems and Signal Processing*, vol. 156, p. 107671, 2021.
- [19] J. Gradišek, A. Baus, E. Govekar, F. Klocke, & I. Grabec, *Automatic chatter detection in grinding*, *International Journal of Machine Tools and Manufacture*, vol. 43, no. 14, pp. 1397–1403, 2003.
- [20] K. Li, S. He, B. Li, H. Liu, X. Mao, & C. Shi, *A novel online chatter detection method in milling process based on multiscale entropy*

- and gradient tree boosting, Mechanical Systems and Signal Processing, vol. 135, p. 106385, 2020.
- [21] D. Aslan & Y. Altintas, *On-line chatter detection in milling using drive motor current commands extracted from cnc*, International Journal of Machine Tools and Manufacture, vol. 132, pp. 64–80, 2018.
- [22] , &LI, Z., SONG, Q., JIN, P., LIU, Z., WANG, B. MA, H. (2023). Chatter suppression techniques in milling processes: A state of the art review. Chinese Journal of Aeronautics.
- [23] Nadi, A., & Moradi, H. (2019). Increasing the views and reducing the depth in random forest. Expert Systems with Applications, 138, 112801.
- [24] Feng, Z., Yi, H., Shu, A., & Tang, L. (2023). Simulation of grinding surface topography considering wheel wear and wheel vibration. The International Journal of Advanced Manufacturing Technology, 130(1–2), 475–490.
- [25] M. C. Yesilli & F. A. Khasawneh, *On transfer learning of traditional frequency and time domain features in turning*, in Volume 2: Manufacturing Processes; Manufacturing Systems; Nano/Micro/Meso Manufacturing; Quality and Reliability, American Society of Mechanical Engineers, Sep 2020.
- [26] Dutta, C., Nagarajan, N., Adhoni, Z. A., Dhandapany, S. K., Alagarsamy, A., Chakravarthy, A. K., Praveen, N., Pushpanathan, G., & Kannan, S. (2023). Outlier detection in additive manufacturing using novel machine learning algorithm. The International Journal of Advanced Manufacturing Technology. <https://doi.org/10.1007/s00170-023-12798-9>
- [27] Jia, W., Sun, M., Lian, J., & Hou, S., *Feature dimensionality reduction: A Review*. Complex Intelligent Systems, 8(3), 2663–2693. 2022.
- [28] Navarro-Devia, J. H., Chen, Y., Dao, D. V., & Li, H. (2023). Chatter detection in milling processes—a review on signal processing and condition classification. The International Journal of Advanced Manufacturing Technology, 125(9–10), 3943–3980.
- [29] Xia, S., & Yang, Y. (2023). A Model-Free Feature Selection Technique of Feature Screening and Random Forest-Based Recursive Feature Elimination. International Journal of Intelligent Systems, 2023, 1–16.
- [30] Wang, Q., Song, L., Zhao, J., Wang, H., Dong, L., Wang, X., & Yang, Q. (2022). Application of the gradient boosting decision tree in the online prediction of Rolling Force in hot rolling. The International Journal of Advanced Manufacturing Technology, 125(1–2), 387–397.
- [31] Christ, M., Braun, N., Neuffer, J., & Kempa-Liehr, A. W., *Time series feature extraction on basis of scalable hypothesis tests (tsfresh – a python package)*. Neurocomputing, 307, 72–77. 2018.

Appendix A Figures

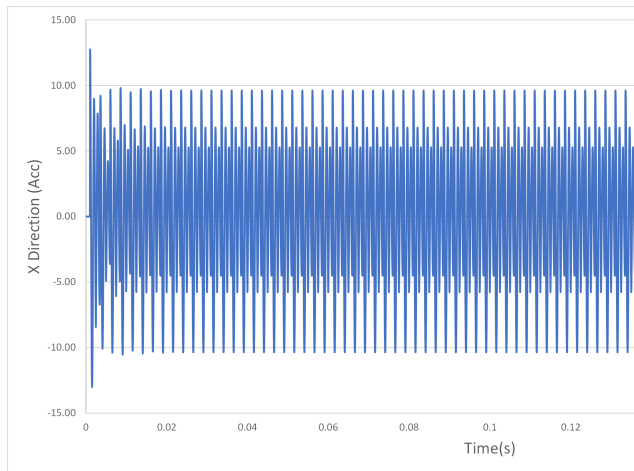


Fig. A1: Example vibration acceleration, stable.

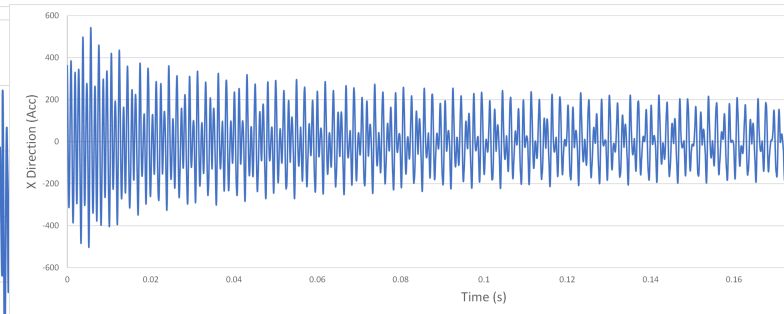


Fig. A4: Example vibration acceleration slightly above chatter line, unstable.

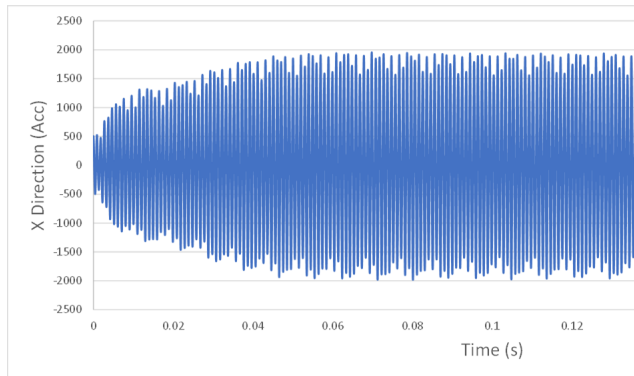


Fig. A2: Example vibration acceleration, unstable.

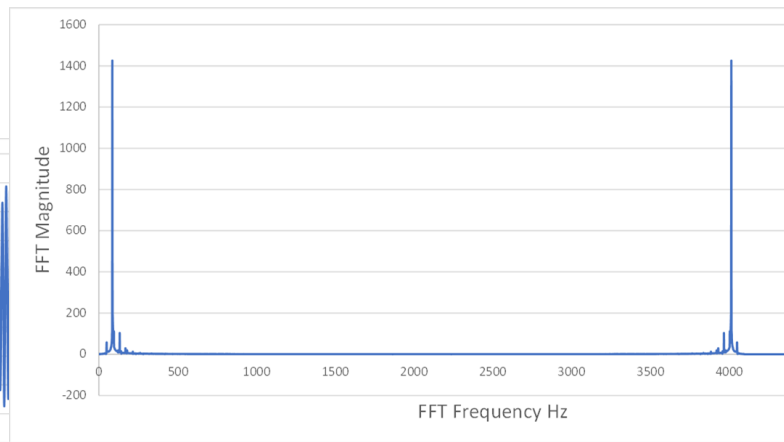


Fig. A5: Example FFT data indicating clear instability.

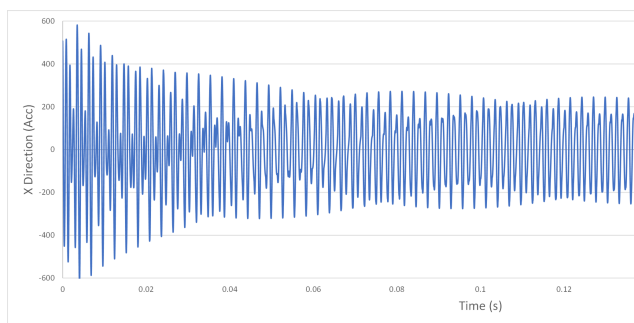


Fig. A3: Example vibration acceleration slightly below chatter line, stable.

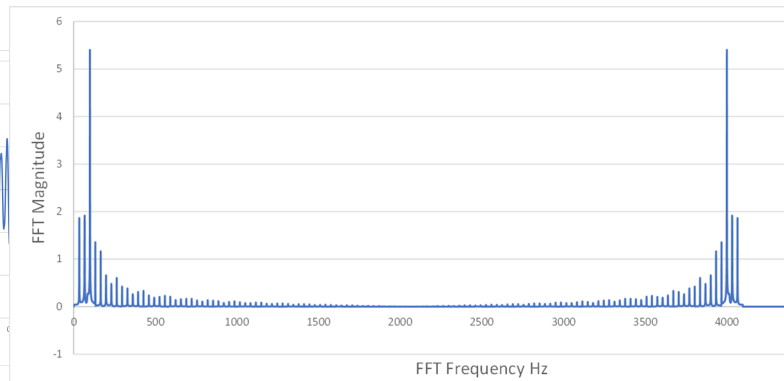


Fig. A6: Example FFT data indicating clear stability.

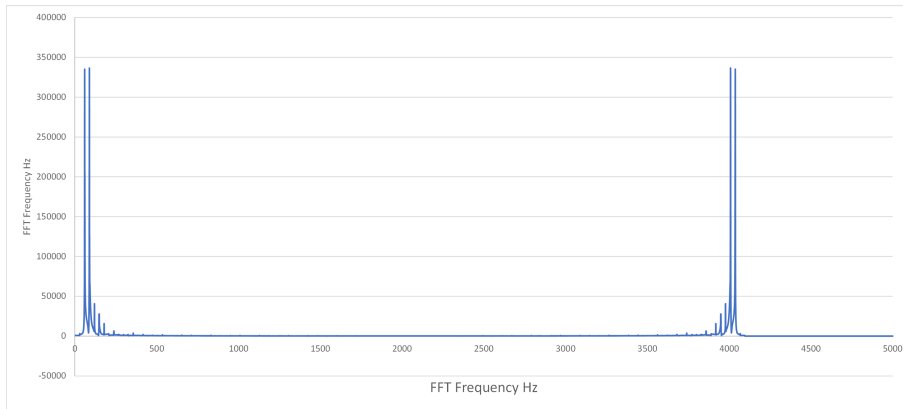


Fig. A7: Example FFT data slightly above chatter line, unstable.

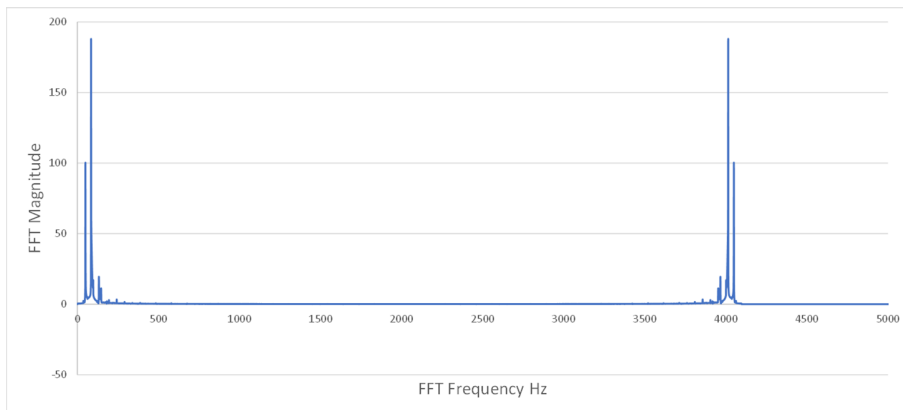


Fig. A8: Example FFT data slightly below chatter line, stable.

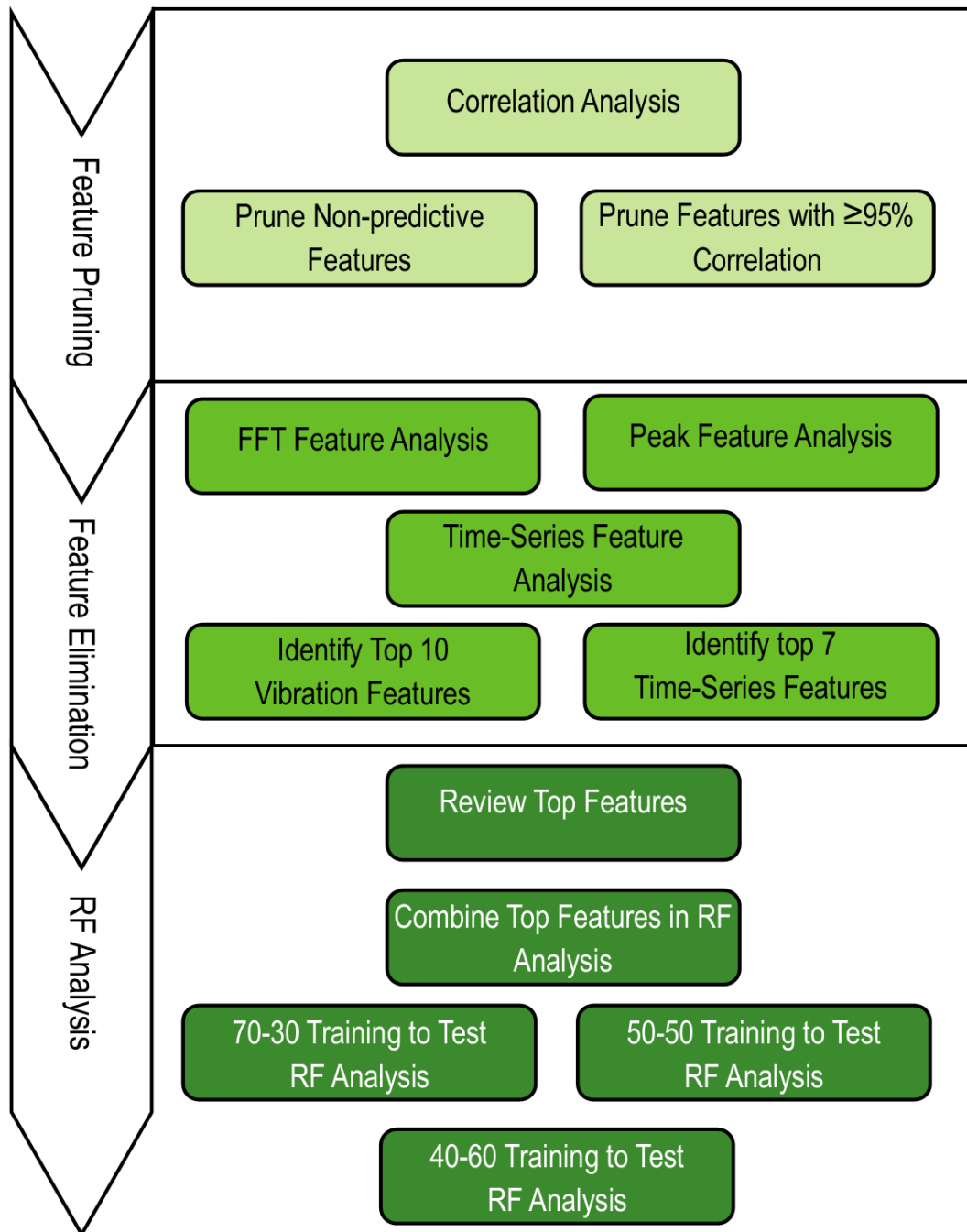


Fig. A9: The research approach includes feature pruning, recursive feature elimination, and RF analysis. Steps in this approach are provided with a generalized flow from top to bottom.