Group no: 13

Ashok Santosh Menon : 220103023
Aditya Singh : 220103003
Amit : 220103010

Chatter Detection in Machining:

1. Introduction

Chatter in machining processes is a self-excited vibration that affects the quality of the surface finish, tool life, and overall machining performance. This project aims to detect chatter using machine learning techniques by analyzing vibration signals collected during turning operations under varying cutting conditions. The goal is to develop a robust model capable of identifying chatter occurrences from acceleration data, using both time and frequency domain features.

2. Objective

- To segment acceleration signals into 1-second windows.
- To extract meaningful features from time and FFT domain.
- To use unsupervised learning (KMeans) for initial labeling.
- To semi-automate the labeling process for supervised learning.
- To train and evaluate multiple machine learning models.
- To build a voting-based ensemble classifier for final prediction.
- To deploy an interactive notebook-based web app for chatter detection.

3. Dataset Description

The dataset consists of 12 CSV files, each corresponding to different spindle speeds and depths of cut. Each file contains acceleration vs. time data for X, Y, and Z axes. Sampling frequency was determined to be approximately 407 Hz.

4. Methodology

4.1 Data Preprocessing

- All CSV files were read and mapped by condition.
- Signals were segmented into 1-second windows.

- Sampling frequency was computed from the time column.

4.2 Feature Extraction

Time Domain Features:

- RMS
- Standard Deviation
- Skewness
- Peak-to-Peak
- Time kurtosis

FFT Domain Features:

- Dominant Frequency
- Band Energy Ratio
- Spectral Kurtosis
- Peak Count
- Spectral Centroid

4.3 Unsupervised Learning (KMeans)

- Combined time and frequency features.
- Applied KMeans clustering to generate initial chatter/stable labels.
- Validated cluster-wise feature distributions and representative plots.

4.4 Semi-Automated Labelling

- Clusters were inspected visually using FFT and time domain plots.
- Based on patterns and prior knowledge, labels were refined.

4.5 Supervised Learning Models

The following models were trained using cross-validation:

Random Forest

An ensemble of decision trees that works well with mixed features. It showed high accuracy and low variance.

XGBoost

A gradient boosting technique optimized for speed and performance. It gave comparable performance to Random Forest with slightly different feature importances.

Support Vector Machine (SVM)

Classifies by finding an optimal separating hyperplane. While highly precise, it showed lower recall for chatter.

K-Nearest Neighbors (KNN)

Classifies based on majority voting among nearest neighbors. It performed well but was sensitive to noise.

5. Model Evaluation

Models were evaluated using cross-validation. Metrics included accuracy, precision, recall, F1-score, and ROC AUC:

Model	Accuracy	Precision	Recall	F1	ROC AUC Mean	ROC AUC Std
XGBoost	0.9911	0.9840	0.9759	0.9798	0.9995	0.0004
Random Forest	0.9929	0.9918	0.9758	0.9837	0.9995	0.0006
SVM	0.9619	1.0000	0.8273	0.9054	0.9971	0.0019
KNN	0.9778	0.9956	0.9038	0.9473	0.9880	0.0071

Feature Importance

- **Random Forest/XGBoost**: RMS was most important.
- **SVM**: Peak Count had the highest influence.
- **KNN**: Spectral Kurtosis played a major role.

6. Ensemble Model

A Voting Classifier ensemble (Random Forest, XGBoost, SVM, KNN) was implemented.

- **Voting = 'soft'**
- Improved prediction robustness.
- Leveraged complementary strengths of individual models.

7. Deployment

A Google Colab-based interactive notebook web app was developed with the following features:

- Upload a CSV file.
- Segment it into 1-second windows.
- Extract time + FFT features.
- Load the trained ensemble model (.pkl).
- Predict chatter/stable per segment.
- Plot FFT and time domain signals for visualization.

8. Key Observations

- Chatter was observed to correlate with higher depths of cut and lower spindle speeds.
- Time domain features like RMS and peak-to-peak were reliable indicators.
- Spectral features like kurtosis and band energy ratio helped highlight unstable segments.
- Ensemble model performed better than any single model in terms of generalization.

9. Conclusion

The developed chatter detection model effectively identifies chatter across varied machining conditions using acceleration data. Time and frequency domain features offer valuable insights into the nature of vibrations. The ensemble classifier ensures reliable classification by aggregating predictions. The interactive app allows practical usage of the model, making it a deployable solution for real-world monitoring in smart manufacturing environments