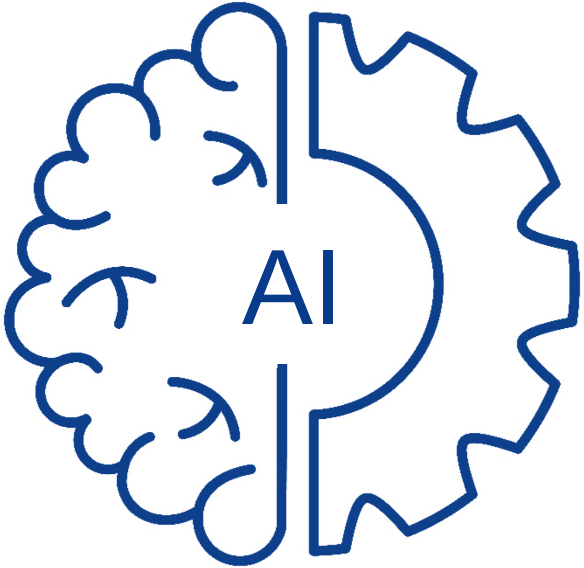
Data Augmentation based Surface Quality Monitoring with Machine Learning Models

Yaoxuan Zhu, PhD Candidate

Overview



- Intelligent Machining System
- Surface Quality Monitoring
- Prediction Model Building
- Summary & Future Work

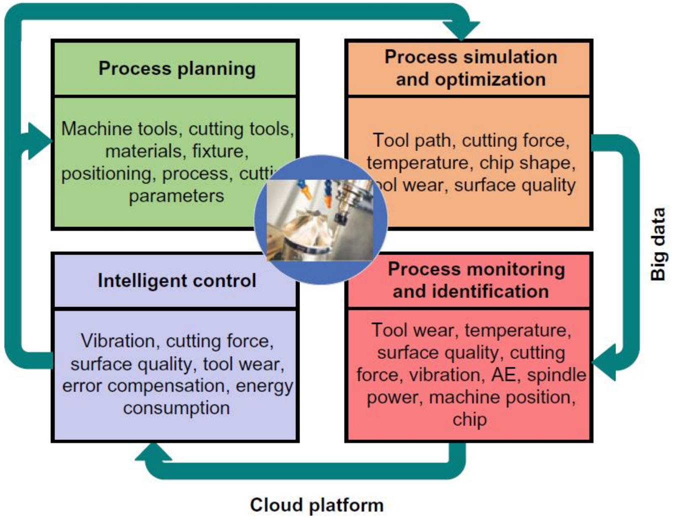
Intelligent machining

Intelligent machining is a process that based on **process monitoring** and **intelligent control technology**, designed to adaptively **solve many uncertain problems** in the process that require manual intervention. Its ultimate goal is to realize the **intelligent decision-making**, **monitoring**, and **control of the machining process**.

Establish the interaction with different systems

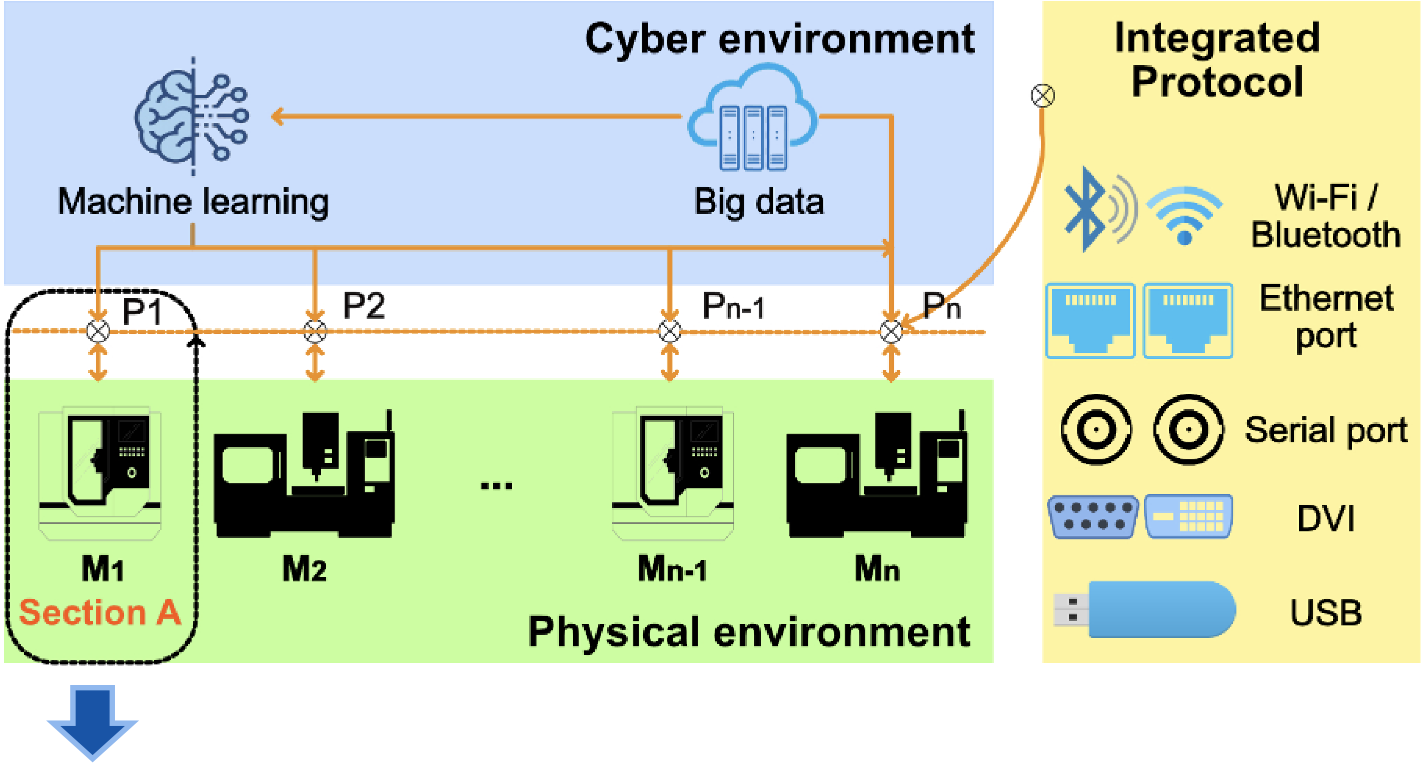
- Machine tools
- Sensors
- Controller network
- Big data
- Cloud-based data storage system

- Monitor and extract features
  - Machine tools
  - Cutting tools
  - **Workpiece Quality (Surface Roughness)**

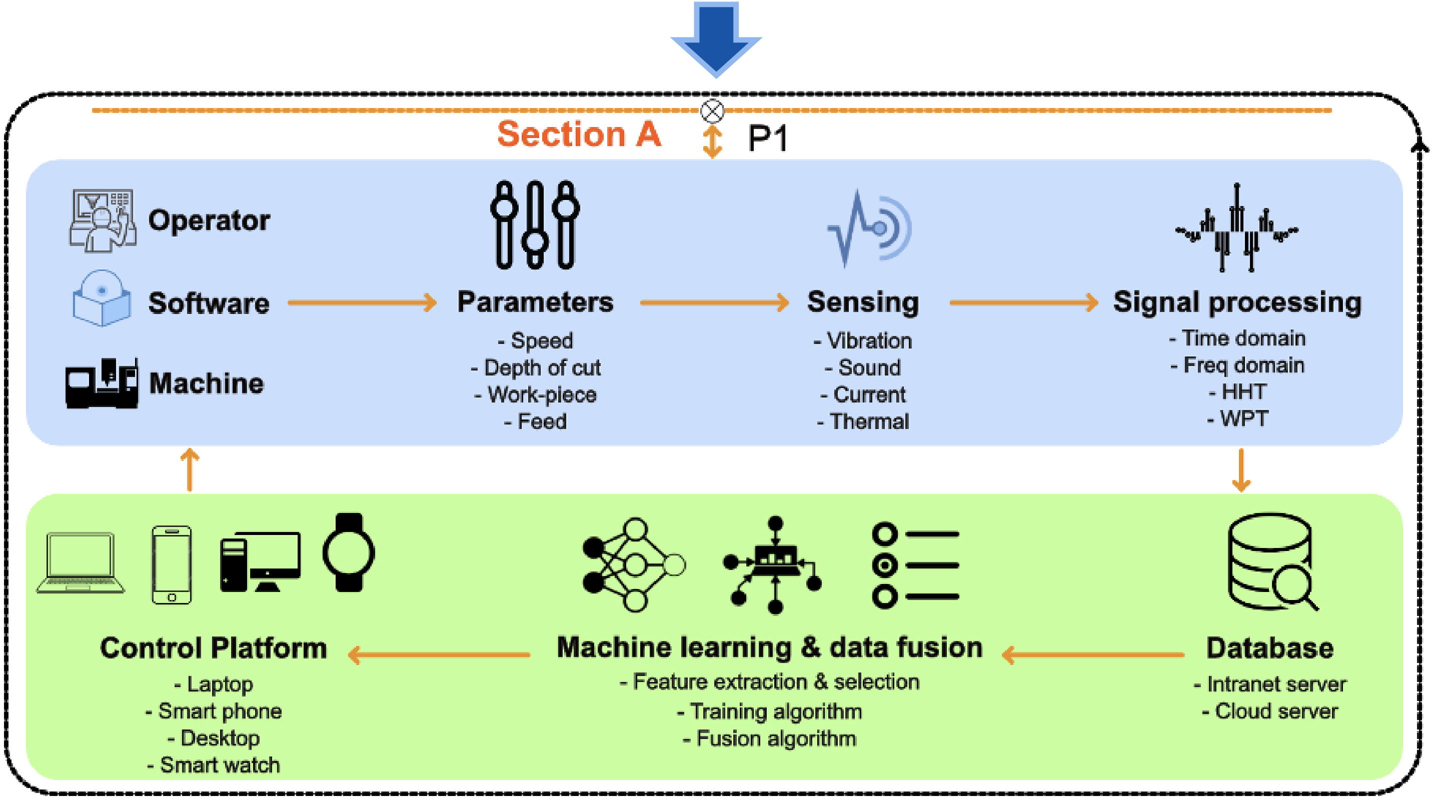


Reference Zhu, Kunpeng. Machining Systems | Zhou Z, Xie S, Chen D (2012) Fundamentals of digital manufacturing science. Springer

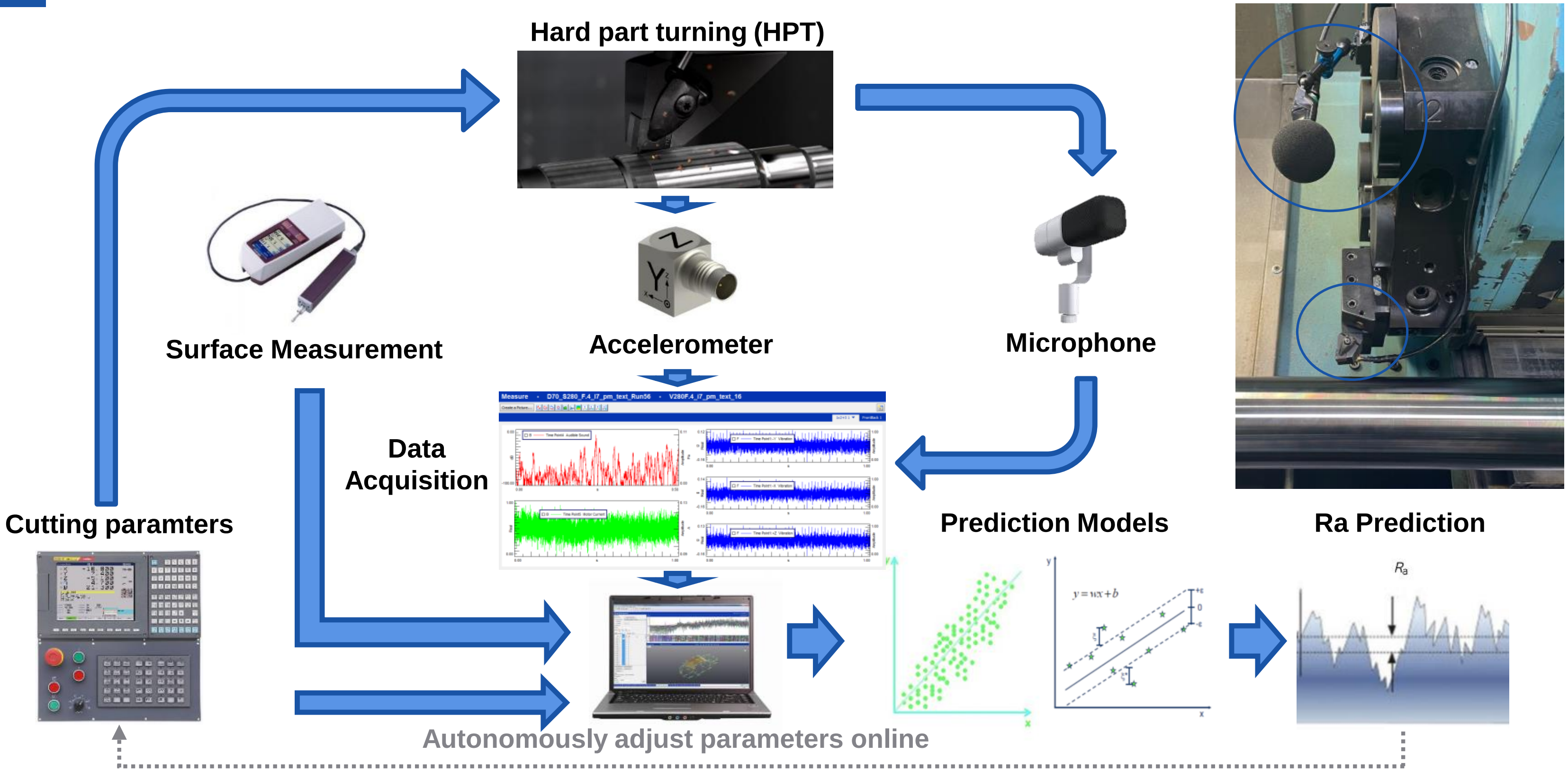
Conceptual diagram for smart machining



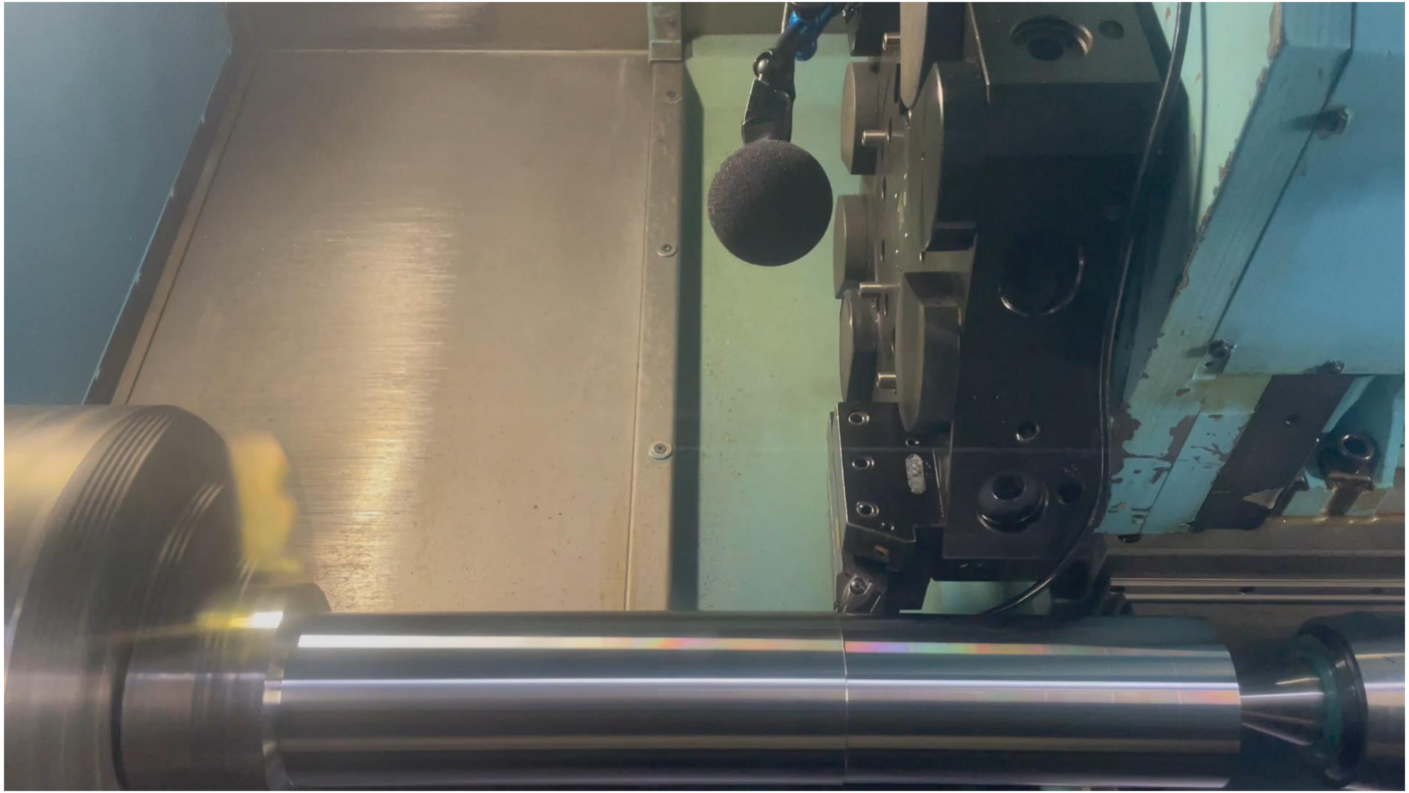
Conceptual diagram for smart machining



Surface Quality Monitoring & Prediction



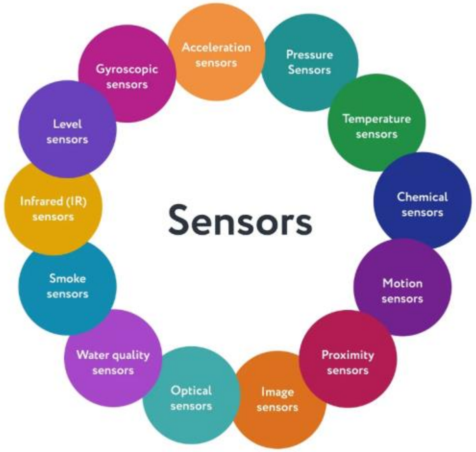
Hard Part Machining with Multi-sensors



Core technologies & challenges

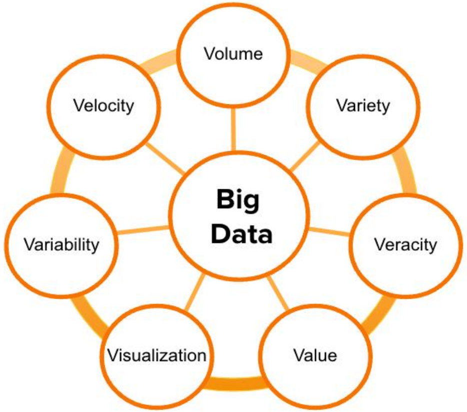
Core technologies

- Sensor integration, sensor network integration
- Integration of communication protocols
- Machine/deep learning algorithm
- Data acquisition of machine tools, process and products

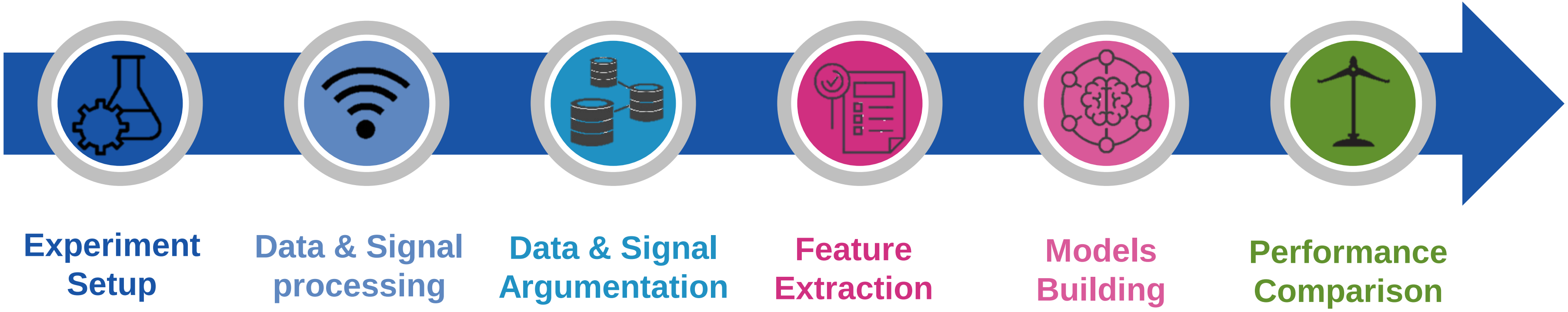


Challenges

- **Big data accessibility**
- **Multi-sensors integration**
- Safety guaranteeing
- Security of smart machining system

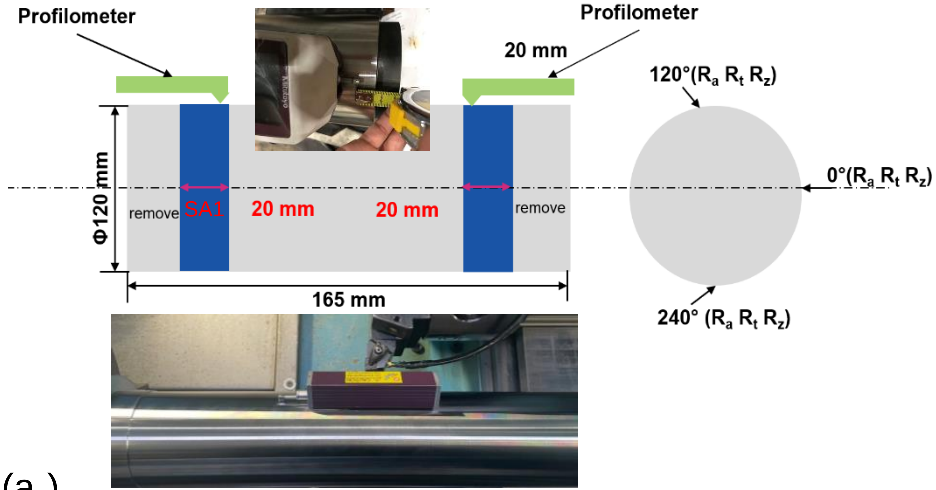


Prediction Model Building



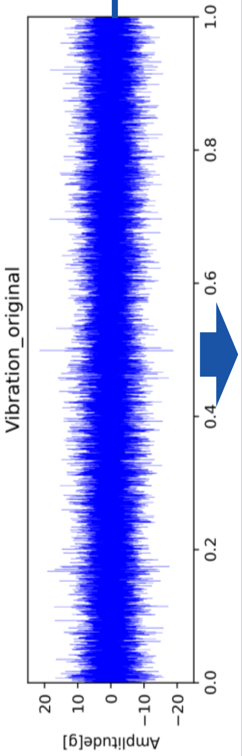
Experiment setup

- Surface roughness measurement
  - ❖Profilometer: Mitutoyo SJ-210
  - ❖Cut-off: 0.8
  - ❖Sampling points: 5
- Vibration measurement
  - ❖Accelerometer: Dytran 3023A2
  - ❖Frequency range: 20 Hz – 10k Hz
  - ❖3-axes: feed (a_f), radial (a_p), tangential(a_c)



V_c (m/min)	a_p (mm)	F (mm/rev)	L (mm)	D (mm)	SS (RPM)	T (s)	CNC Machine
150	0.1	0.15	170	120	398	170	Swedturn (KTH)
270	0.1	0.15	170	120	716	94	Swedturn (KTH)

Signal Processing & Augmentation



Gaussian Noise

$$\tilde{X} = X + \alpha_{gaus} G$$
$$G \sim N(0,1)$$

Masking Noise

$$\tilde{X} = X + \alpha_{mask} G$$

Signal Translation

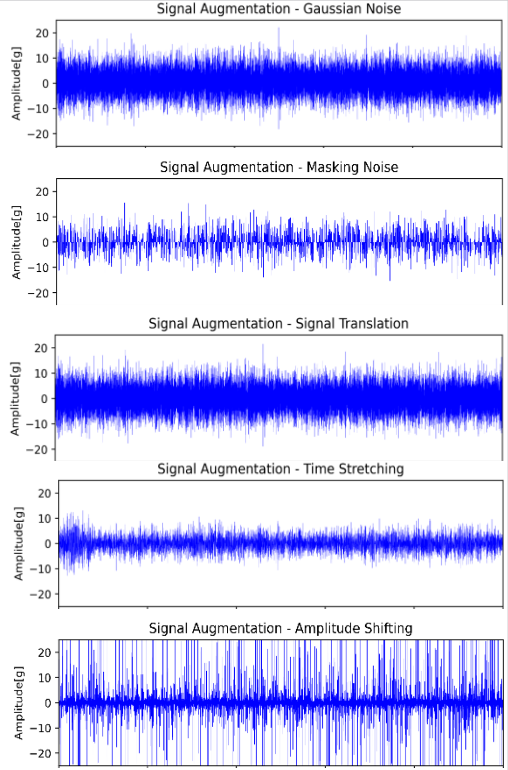
$$\tilde{X} \in [-\alpha_{tran}, \alpha_{trans}]$$

Time Shifting

$$\tilde{X} \in [1, \alpha_{stre}], [\alpha_{stre}, 1]$$

Amplitude Shifting

$$\tilde{X} = \alpha_{scale} X$$



Mean

$$\mu = \frac{\sum_{i=1}^{N} xi}{N}$$

Standard Deviation

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}}$$

Peak to Peak

$$pp = X\_max - X\_min$$

Entropy

$$H = -\sum_{i=1}^{N} P(x_i) log P(x_i)$$

Skewness

$$Sk = \frac{\sum_{i=1}^{N}(x_i - \mu)^3}{(N-1)\sigma^3}$$

Kurtosis

$$Ku = \frac{\sum_{i=1}^{N}(x_i - \mu)^4}{(N-1)\sigma^4}$$

Training
- Original
- Augmented

Testing
- Original

ML Algorithms

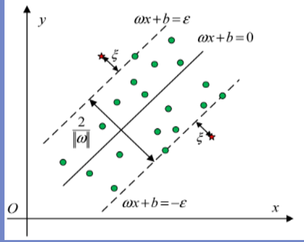
Prediction Models

Support Vector Machine - SVR

- Solve non-linear problem
- Applied with different kernel function

$$\hat{y} = f(x) = \omega^T x + b,$$

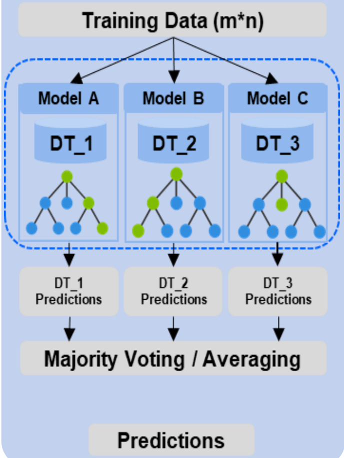
$$\begin{cases} \min \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^{n} (\xi_i + \xi_i^*) \\ s.t. y_i - f(x_i) \le \varepsilon + \xi_i \\ f(x_i) - y_i \le \varepsilon + \xi_i \\ \xi_i \ge 0, \xi_i^* \ge 0 \end{cases}$$



(Lu et al. 2020)

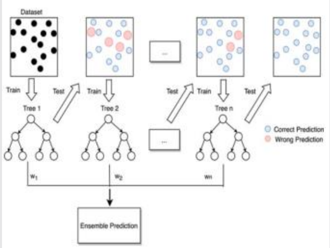
Random Forest - RF

- A tree-based ensemble method
- weak learners
- Use a bootstrap replica to train model



Gradient Boosting Regression - GBR

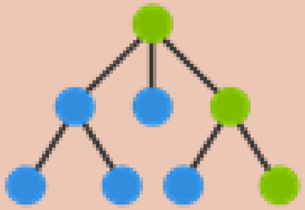
- A tree-based ensemble method
- Suppress potential outlier
- Fit complex nonlinear relationship
- Approximate the true value by the loss function minimization



(Zhang et al. 2021

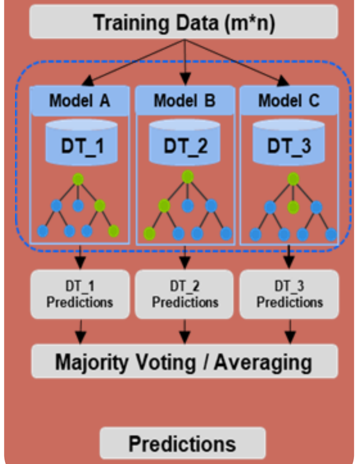
Decision Tree Regression - DTR

- Easy to interpret
- Transparent structure
- Each leaf node of trees is a simple regression model
- Pruning supports to reduce its complexity and improve its robustness



Extra Trees - ET

- An extension of RF
- Avoid overfitting
- Randomly select feature
- Use whole training data



Model Performance Metrics

MAPE

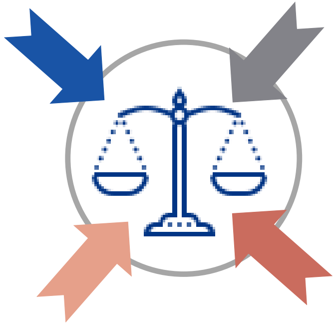
Mean Absolute Percentage Error

$$\bar{e}_r = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{(Ra_i^{Exp} - Ra_i^{Pred})}{Ra_i^{Exp}}\right| \times 100$$

RMSE

Root Mean Square Error

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Ra_i^{Exp} - Ra_i^{Pred})^2}$$



MAE

Mean Absolute Error

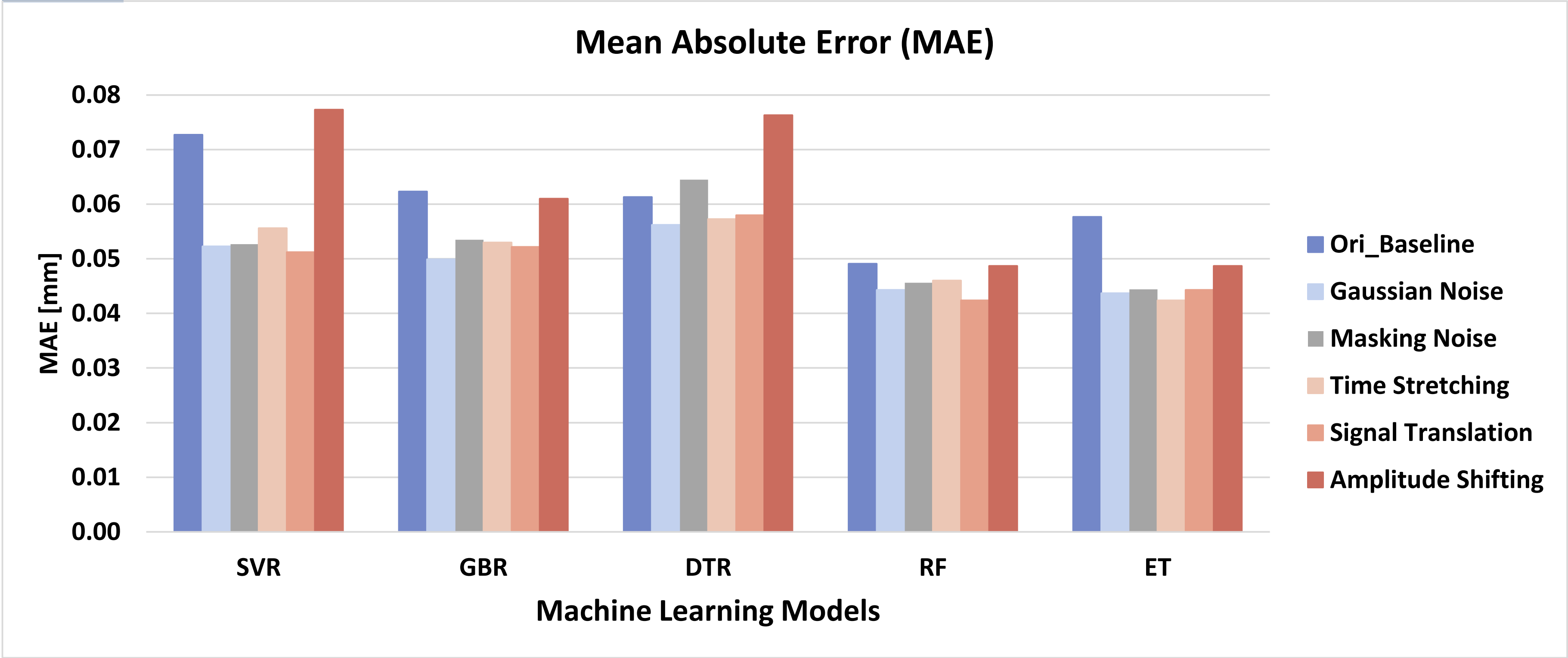
$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|Ra_i^{Exp} - Ra_i^{Pred}\right|$$

R^2

Coefficient of Determination

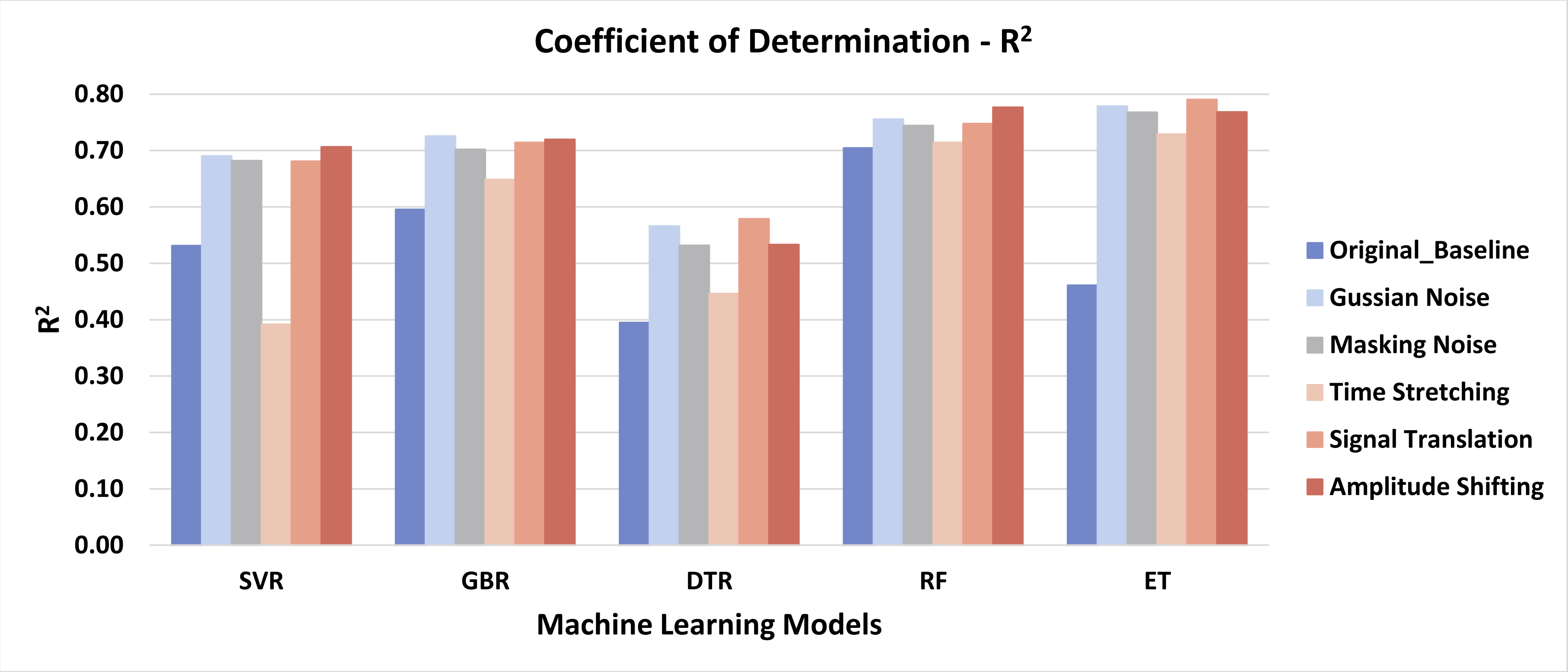
$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Ra_i^{Pred} - Ra_i^{Exp})}{\sum_{i=1}^{n}(Ra_i^{Exp} - \overline{Ra_i^{Exp}})^2}$$

Model Performance Comparison



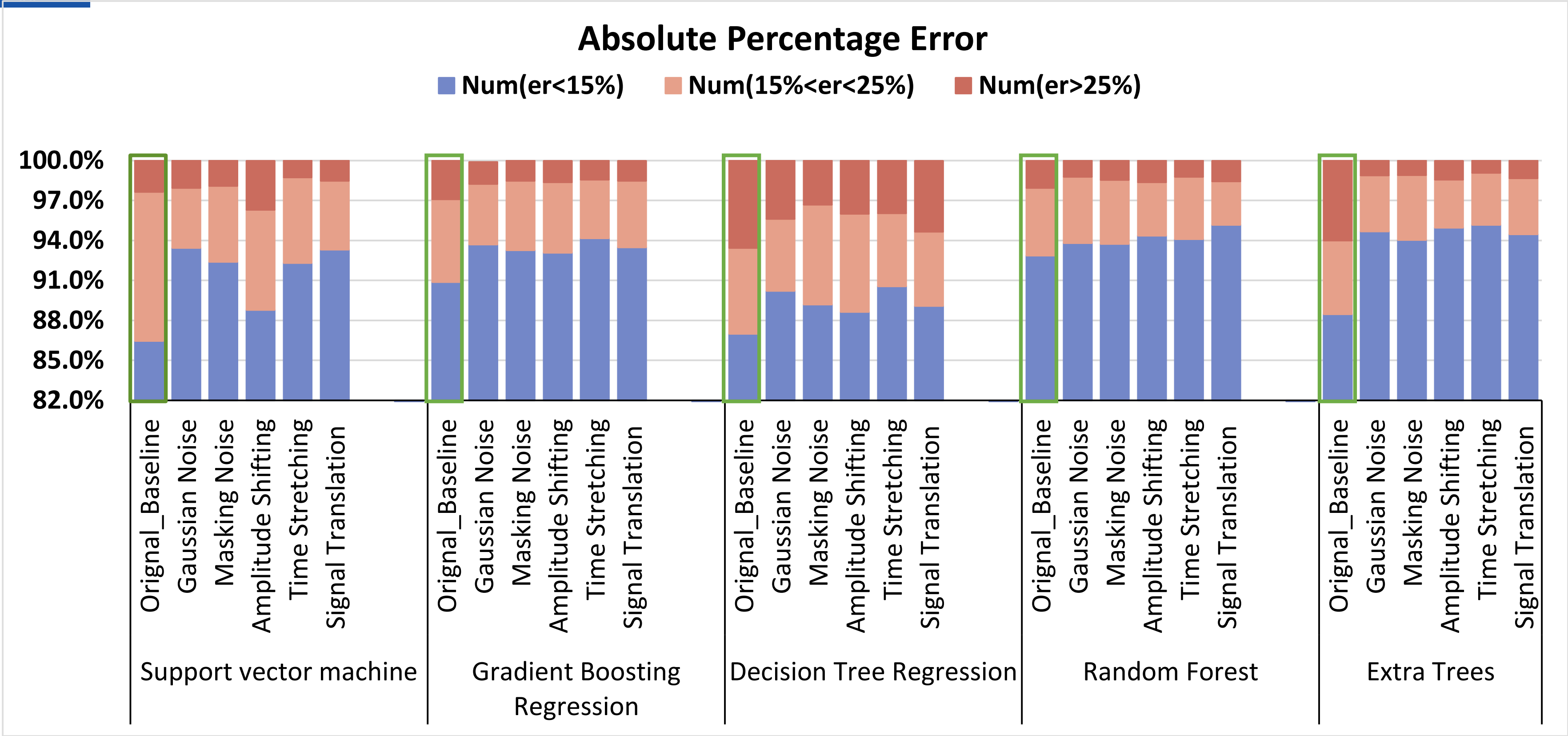
Support vector machine | Gradient Boosting Regression | Decision Tree Regression | Random Forest | Extra Trees

Model Performance Comparison



Support vector machine Gradient Boosting Regression Decision Tree Regression Random Forest Extra Trees

Model Performance Comparison



Summary & Future Work

- ✓ Time series data augmentation is efficient method on enlarging amount of data
  - ➢ Prediction error reduction – lower MAE, higher number of er(<15%)
  - ➢ More robust or generalized – higher R^2
- ✓ All data augmentation methods enable to improve generalization of each algorithm
- ✓ In regard with prediction error reduction, Gaussian Noise outperforms others in all algorithms
- ✓ Extra Trees and Random Forest bring the best outcomes than other algorithms

- o Multi-sensors (features) fusion
- o Spectrogram and generative AI based augmentation method - GANs
- o Advanced deep learning based-algorithm to deploy prediction model