

アンサンブル学習と KDE を用いた生産改善システムによる 管理者支援

Manager support by production improvement system using ensemble learning and KDE

† 赤羽勇飛
Yuhi AKAHANE

† 上条康佑
Kosuke KAMIJO

† 田邊 造
Nari TANABE

公立諏訪東京理科大学 † Suwa University of Science

E-mail:†{T118002@ed, T119045@ed, nari@rs}.sus.ac.jp

1 はじめに

本論文は、アンサンブル学習と KDE を用いたライン式生産改善システムによる管理者支援を提案する。提案手法は、(Step 1) 工場機械に設置した自作 IoT デバイスのセンサデータを Zigbee 通信により工場サーバへ送信する。(Step 2) 工場サーバからインターネットを介して送られてきたデータをメインサーバで解析し、結果をデータベースに蓄積する。(Step 3) 蓄積したデータからアンサンブル学習により生産数の予測と KDE により最適な生産遅延原因を明らかにしている。提案手法の特徴は、(i) 生産数の推定と (ii) 遅延原因工程の特定により生産工程の改善箇所の提案が可能なことである。

2 提案手法

2.1 システム構成

2.1.1 工場サーバ部

図 1(a) 独自設計した IoT デバイスは、工場機械から取得したアナログ/デジタルのセンサデータを Zigbee 通信を用いて、図 1(b) 工場サーバ (Raspberry Pi) に送信している。工場サーバは、IoT デバイスから受信したセンサデータを Node.js によって解析した後に、解析データを MongoDB に蓄積すると共に、インターネットを介して図 1(c) メインサーバのデータベース (MySQL) へ送信する。

2.1.2 メインサーバ部

図 1(c) メインサーバは、データベースに保存されているセンサデータを Python で解析し、Django を用いて MVT モデルの Web アプリケーションを作成することで、図 1(d) クライアントに解析ページを提供している。

2.2 アンサンブル学習

2.2.1 学習方法

提案手法は、図 2 の手順で生産数の推定を回帰問題としてアンサンブル学習を用いて解決している。学習器は XGBoost, LightGBM および CatBoost を用い、学習モデルの評価関数は RMSE (Root mean square error)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - f_i)^2} \quad (1)$$

としている。ここで、データ数 n 、真値 y_i 、予測値 f_i であり、データは 5 回シード値を変えながら KFold で 5 分割することでモデルの汎化性能を高める。なお、アンサンブルは平均法を用いた。

2.2.2 生産数の推定

生産数の推定は、説明変数と目的変数を各生産工程の平均サイクルタイムと 1 日の生産数としている。これにより、サイクルタイムが分かった時点で生産数の推定が可能となる。

2.3 ノンパラメトリック推定

遅延原因工程の推定は、記録したサイクルタイムの正常データから KDE (Kernel Density Estimation) により確率密度分布を推定し、分布を異常度にする事でサイクルタイムの異常検知をしている。

データ X_i とバンド幅 h およびカーネル関数 K としたとき

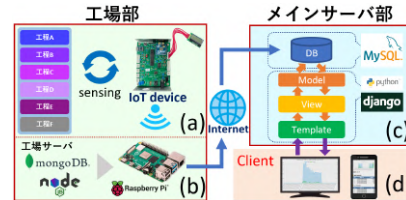


図 1: システム構成

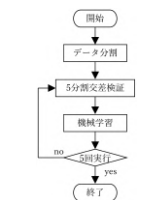


図 2: 学習手順

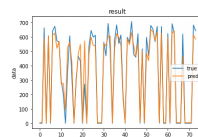


図 3: 予測結果

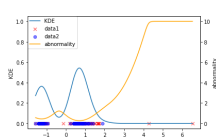


図 4: KDE 推定結果

表 1: RMSE 比較

| 学習手法 | RMSE |
|----------|-------|
| XGBoost | 128.5 |
| LightGBM | 123.3 |
| CatBoost | 116.2 |
| ensemble | 95.4 |

推定分布 $p(x)$ は次式で求められる。

$$p(x) = \frac{1}{Nh} \sum_{i=1}^N K\left(\frac{x - X_i}{h}\right) \quad (2)$$

ここで、カーネル関数 K は Gaussian Kernel を用いた。バンド幅 h は LSCV (Least Squares Cross Validation) で推定している。

分布の逆数に対数をとった値を異常度とし、異常データと異常度の分布から閾値を決め、一定の異常度を越えた場合を遅延原因工程 (異常) として判断している。

3 シミュレーション

3.1 生産数の推定

図 3 は青が真値、橙がアンサンブル学習の結果により求められた生産数の推定値を示している。図 3 の結果から、おおよそ適切な生産数の推定が可能となった。次に、表 1 は学習モデルの評価のためにアンサンブルに用いた学習器とアンサンブル学習の RMSE を示している。表 1 より、アンサンブル学習によって各学習器の個別での結果より精度が高くなった。

3.2 遅延原因工程の推定

図 4 はライン内の 1 工程に推定された確率密度分布、異常度およびデータを示している。データに対しておおよそ適切な分布が形成できたため異常度の推定が可能となった。これにより、サイクルタイムが異常度 1 を超えるかどうかで遅延原因工程であるかを判断可能となる。

4 まとめ

本論文は、アンサンブル学習と KDE を用いたライン式生産改善システムによる管理者支援を提案した。提案手法は、生産数が目標生産数に到達できるか事前に把握可能となり、遅延原因工程の推定から生産管理者の支援が可能となった。

参考文献

- [1] 白鳥剛郎, 三石大, 吉廣卓哉, “データベースビッグデータ時代の基礎—”, 共立出版, Sep. 2014.
- [2] 坂本俊之, “作ってわかる! アンサンブル学習アルゴリズム入門”, 株式会社 C&R 研究所, Jun. 2019.