

ニューラルネットワークを用いたガス製造プラントの品質予測

Predicting Quality of Gas Product Using Neural Networks

泉谷知範 切通恵介 島田健一郎 伊藤浩二
Tomonori Izumitani Keisuke Kiritoshi Ken'ichiro Shimada Koji Ito

NTT コミュニケーションズ株式会社
NTT Communications Corporation

IoT, AI 技術の発展に伴い, 化学プラントにおいても, 温度, 圧力などのプロセスデータを用いた異常の検知, 品質の予測への期待が高まっている. 本研究では, 実際のガス製造プラントを対象とし, 51 種のプロセスデータから, 製品品質のインジケータとなる X ガス濃度を予測する技術を提案する. ニューラルネットワークを利用することで, 20 分後の X ガス濃度を線形回帰など従来手法に比べ高精度に予測できることを確認した.

1. はじめに

近年, センサやネットワークに関する技術の向上, データ収集環境の整備などにより, IoT(Internet of Things) データの活用が拡大している. 特に, 製造業などの工業プロセス, インフラなどのシステム管理, プロセス制御などを統合的に扱う SCADA(Supervisory Control And Data Acquisition) と呼ばれるシステムの普及などにより, 大量のデータが収集され, センサデータや制御データなどの時系列データ (プロセスデータと呼ばれる) が, 異常の監視などプロセスの管理に利用されている.

多くの工業システムでは, プロセスデータに定められた許容範囲から逸脱が生じた場合アラートを上げたり, 熟練の運転員がプロセスデータを表示したチャートを確認することで異常を監視しているが, 複数のプロセスデータが関連した複雑な異常検出ルールを決定することが困難であることや, 現在起きている異常のみしか検知できないため未然の対応が難しいという問題があった.

本研究では, このような問題を解決するために, プロセスデータを用いて, 監視対象のセンサ値など特定プロセスデータの未来の値を予測する問題を取り扱う.

従来より, 測定に高いコストのかかる等の理由により常時計測できないセンサの値を, 測定可能なセンサ値から推定する技術として, ソフトセンサ技術が利用されて来た [船津 2014]. ソフトセンサは, 物理または化学的な理論をもとに専門家が計算モデルを設計するのではなく, 収集された多量のデータから, 複数のセンサデータを入力とし, 特定のセンサ値を推定する, 機械学習の回帰問題として定式化される. 未来の異常などの予測についてはこれまではあまり検討されていなかったが, 機械学習による実現が期待されている [Reis 2014].

一方, 回帰, 分類などの機械学習の分野では, 特に画像や音声の認識などのタスクにおいて, ディープニューラルネットワークの利用が急速に広まっている. 学習アルゴリズムの進展や GPU などの並列計算環境の向上により, 深く, 複雑なネットワークが学習可能となって来ていることから, 従来主流であった特徴抽出とサポートベクトルマシン (SVM) などの機械学習手法の組み合わせに比べ高い認識精度を実現している [Goodfellow 2016].

ソフトセンサをディープニューラルネットワークで構成する研究例はまだ少ないが, 例えば Shang らは, 石油精製プラントにおけるソフトセンサを, 4 層の Deep Belief Networks で構成し, SVM, PLS (Partial least square) などの従来的手法を上回る精度を実現している [Shang 2014]. また, Qui らは, 下水処理場における生化学的酸素必要量 (BOD) の推定に Stacked Auto-Encoder(SAE) を用いた [Qiu 2015].

本研究では, ガスを製造する化学プラントにおいて実際に収集されたプロセスデータを入力として, 製造されるガスの品質のインジケータとなる不純物 (仮に X ガスとする) 濃度の未来の値を, ニューラルネットワークを用いて予測する方法を提案する. X ガス濃度の予測により, 将来の品質の悪化を事前に予測し早期の対処を行うことや, 実際の X ガス濃度分析計と推定値を二重化させることにより保全の改善が期待される.

2. 方法

対象とする化学プラント (三井化学株式会社) では, 配管を經由して流入させた原料を反応炉において高温で反応させることにより製品となるガスを製造している. 基本的に反応炉の中では常に同じ化学反応を起こしており, 原料等が連続的に投入され, 反応物であるガスも連続的に反応炉から排出される仕組みとなっている. このようなプロセスは連続プロセスと呼ばれる.

製造されるガスの品質は X ガス濃度測定装置により常時測定されており, この値やその他のプロセスデータを監視することにより品質の管理を行っている. 本研究においては, 異常に

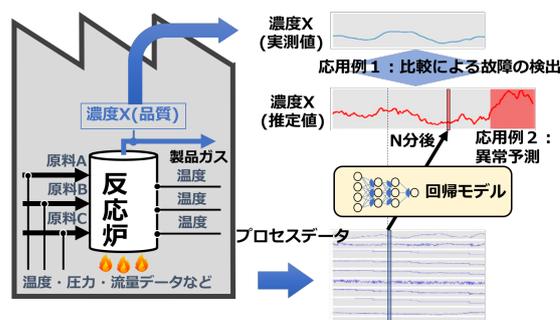


図 1: 予測システムの概要

連絡先: 泉谷知範, NTT コミュニケーションズ株式会社, 東京都港区芝浦 3-4-1 グランパークタワー 16F, 050-3812-4686, tomonori.izumitani@ntt.com

対する早期の対応などを目指し、未来の X ガス濃度の予測の実現性を評価する。

図 1 は、本研究で提案する X ガス濃度の予測を行うシステムの概要を示している。反応炉、配管などには温度、圧力、流量などを測定するセンサが設置されており、計 51 種類のプロセスデータを 1 分ごとに収集している。また、品質のインジケータとなる X ガスの濃度も同様に 1 分ごとに収集している。

プロセスデータは各時刻において抽出され、ベクトル値として表現される。このベクトル値を入力とし、 N 分後の未来の X ガス濃度を出力とするニューラルネットワークを構成する。ニューラルネットワークの学習には、説明変数に各時刻のプロセスデータ、目的変数に N 分後の X ガス濃度の値を用いる。

前処理として、プロセスデータごとに学習データにおける平均と標準偏差を用いた正規化を行った。また、ほとんど一定値を示すプロセスデータ 6 種類と計器が故障していることがわかっているデータ 1 種類を除外した 44 種類のプロセスデータを解析対象として用いた。

目的変数である X ガス濃度は、その性質上全ての値が 0-1 の範囲に限定される。この値を目的変数に使うと、予測時に範囲を超えてしまうという問題が発生するため、本研究では、X ガス濃度をロジット変換した値を説明変数として実際には用い、予測する場合はその逆関数である標準ロジスティック関数(シグモイド関数)を適用した値を実際の予測値として用いた。

X ガス濃度予測結果の評価は、実測値とのピアソンの積率相関係数および平均二乗誤差の平方根 (RMSE) を用いて行った。

3. 結果と考察

実験に使用したプロセスデータおよび X ガス濃度は、ある年の 7 月 1 日 0:00 から 10 月 26 日 0:00 に渡り 1 分毎に収集されたものである。プロセスデータのの一つ一つは X ガス濃度との相関は低く、51 種類あるプロセスデータの中で、X ガス濃度との相関係数が絶対値 0.5 を超えたものは 1 つ (相関係数 = 0.50) のみであった。

実験では、中間層 2 層 (それぞれ 100 ノード) 多層パーセプトロン (MLP) を用いた。活性化関数 tanh, 最適化手法はモメンタム項ありの SGD を用いた。比較手法としては重みに対して L1 正則化を行う重回帰分析手法である Lasso 法を用いた。回帰係数へのペナルティは複数試し、一番精度の高かった 0.001 を採用した。

機械学習による回帰、分類問題では、サンプルをランダムにシャッフルしてデータを分割して教師データとテストデータの組み合わせを作るクロスバリデーションで評価が行われるが、本研究においては、実運用を考慮して、オーバーラップを許して全体を 6 つの期間 (約 40 日間ずつ) に分割し、最初の 1ヶ月分のデータを用いて学習し、続く 10 日間のデータを用いて相関係数と RMSE の評価を行った。これは、プラントの状態は経年変化が大きいため、未来のデータまで学習に含めしまうと公正な評価とはならないためである。

表 1, 2 に 20 分後 ($N = 20$) の X ガス濃度を予測する実験の結果を示す。各期間における精度を RMSE と相関係数で表している。表 1 における RMSE の単位は %FS (Full scale) である。RMSE, 相関係数ともに平均して MLP が Lasso 法を上回る結果が得られた。特に RMSE においては平均で 0.5% FS の差が出た。期間 5 においては、ともに精度が低くなっているが、学習に用いた期間とテストに用いた期間で X ガス濃度値の範囲が異なっている (学習時は主に 0.1-0.3 の範囲に収まっているが、テスト時には 0.1-0.6 の範囲にばらついている)

表 1: 期間ごとの予測精度の比較 (RMSE: 単位%FS)

期間	1	2	3	4	5	6	平均
Lasso	1.8	4.5	2.2	2.4	5.7	2.3	3.2
MLP	1.7	4.1	2.2	2.6	3.6	2.1	2.7

表 2: 期間ごとの予測精度の比較 (相関係数)

期間	1	2	3	4	5	6	平均
Lasso	0.90	0.83	0.88	0.80	0.76	0.92	0.85
MLP	0.89	0.83	0.88	0.78	0.87	0.93	0.86

ことが精度低下の原因と考えられる。このような場合においてもニューラルネットワークを用いることで、RMSE が 5.7% から 3.6%, 相関係数が 0.76 から 0.87 へと改善していることから、学習時には未知の範囲についてもある程度モデル化ができていていることを示している。

また、 N の値を変えた実験 (Lasso) では、RMSE で $N = 30$ 分、相関係数で $N = 20$ 分において精度のピークがあり、それを超えると急激に精度が落ちるという傾向があった。プラントの現場への聞き取りにより、プロセスの特性 (遅れ) を考慮すると現場の感覚とも一致するとの見解を得た。

4. おわりに

本研究では、化学プラントのプロセスデータを用いて、製造物の未来の品質を予測する方法を提案した。実際の化学プラントのデータによる実験で、ニューラルネットワークを用いることで、RMSE 2.7%, 相関係数 0.86 と高い精度で予測できることを確認した。近年画像認識等で成功を取めている更に深いネットワークによる精度の向上や、異常が予測された時の要因の抽出などが次の課題となっている。

謝辞

本研究を進めるにあたり、課題の特定、プロセスデータの提供など、ご助力をいただきました三井化学株式会社様に深く感謝いたします。

参考文献

- [船津 2014] 船津公人, 金子弘昌, (2014). ソフトセンサー入門, コロナ社.
- [Goodfellow 2016] Goodfellow, I., Y. Bengio, and A. Courville, (2016). Deep Learning, MIT Press.
- [Qiu 2015] Qiu, Y., Y. Liu, and D. Huang, (2016). Date-Driven Soft-Sensor Design for Biological Wastewater Treatment Using Deep Neural Networks and Genetic Algorithms. Journal of Chemical Engineering of Japan, 49(10), 925-936.
- [Reis 2014] Reis, M. S., R. D. Braatz, and L. H. Chiang, (2016). Big Data: Challenges and Future Research Directions. CEP Magazine(AIChE), 2016 March, pp. 46-50.
- [Shang 2014] Shang, C., F. Yang, D. Huang, and W. Lyu, (2014). Data-driven soft sensor development based on deep learning technique. Journal of Process Control, 24(3), 223-233.