

Title	Application of Data Mining Tools Based on Artificial Neural Network to Biological Process Operations
Author(s)	黄, 濟華
Citation	大阪大学, 2003, 博士論文
Version Type	
URL	<a href="https://hdl.handle.net/11094/44225">https://hdl.handle.net/11094/44225</a>
rights	
Note	著者からインターネット公開の許諾が得られていないため、論文の要旨のみを公開しています。全文のご利用をご希望の場合は、 <a href="https://www.library.osaka-u.ac.jp/thesis/#closed">〈a href="https://www.library.osaka-u.ac.jp/thesis/#closed"〉</a> 大阪大学の博士論文について <a href="https://www.library.osaka-u.ac.jp/thesis/#closed">〈/a〉</a> をご参照ください。

***Osaka University Knowledge Archive : OUKA***

<https://ir.library.osaka-u.ac.jp/>

Osaka University

氏名	黄 濟 華
博士の専攻分野の名称	博士(工学)
学位記番号	第 17806 号
学位授与年月日	平成 15 年 3 月 25 日
学位授与の要件	学位規則第 4 条第 1 項該当 工学研究科応用生物工学専攻
学位論文名	Application of Data Mining Tools Based on Artificial Neural Network to Biological Process Operations (ニューラルネットワークに基づくデータマイニング手法によるバイオプロセス運転管理)
論文審査委員	(主査) 教授 塩谷 捨明  (副査) 教授 吉田 敏臣    教授 室岡 義勝    教授 原島 俊 教授 卜部 格    教授 小林 昭雄    教授 金谷 茂則 助教授 岸本 通雅

### 論文内容の要旨

本論文は、バイオプロセスにおける遺伝子発現から工業用発酵プロセスに至る種々のデータベースを用いて、プロセス操作法や運転管理のためのニューラルネットワークパターン認識手法の開発について述べたものである。実験室規模の培養データベース、工業生産規模の操業データベース、遺伝子発現プロファイルデータベースを具体的なケーススタディとして用い、それぞれの情報抽出システム及びその利用方法を確立した。本論文は要約と結論を含め、5章からなっている。

1章は、序章である。バイオプロセス運転のために、データベースから有用な情報を抽出する手法、特にニューラルネットワークの応用において本研究の背景をなす知見を概説し、本論文の目的と意義について述べた。

第2章では、自己想起型ニューラルネットワーク (AANN) により抗生物質 *virginiamycin* 生産プロセスのオンライン異常検知システムの構築について述べた。ウェーブレットフィルターバンク (WFB) による素データの圧縮、ノイズの除去、また菌体増殖の特徴を利用した入力データの規格化法、さらに中間ノード出力における正常データセットからの距離を用いた AANN の出力データ評価法を開発し有効性を実証した。

第3章では、AANN と主成分分析 (PCA) による工業発酵プロセスの性能評価の自動決定法を開発した。第1章で得られた結果を踏まえ、データ前処理の段階で平均仮説検定 (MHT) を用いてデータベースから性能評価の識別力があるプロセス変数群を抽出した。MHT による解析の結果、識別能力の低い変数を除き、最も識別力のある変数を用いると多変量解析の総合性能が向上できることを実証した。

第4章では、膨大な遺伝子発現プロファイルデータベースから、自己組織化マップ (SOM) による類似性の探索により数十個の代表パターンを抽出する方法を開発した。さらに、SOM により抽出した発現パターン間の関係、関連性を解析するために、ニューラルネットワーク (ANN) を用い、学習誤差による発現パターン間の関連性の評価方法を探索し、発現パターンの相互関係を抽出するモデルを構築した。

最終章において本論文の内容を要約し、結論を述べた。

## 論文審査の結果の要旨

近年の生物工学の発展とコンピュータの性能の向上により、生物プロセスに用いられる細胞機能の解明から工業発酵プロセス運転に至る様々な段階や局面において大量の生物情報の取得と蓄積が可能となりつつある。物質生産を目指したプロセス運転において、これら蓄積データを有効利用することは今後益々重要になってくるものと考えられる。本研究では生物プロセス運転のために多変量解析やニューラルネットワークなどの手法を用いて、実験室規模の発酵プロセスデータ、工業用発酵プロセスデータ、網羅的遺伝子発現解析データベースから有効な情報を抽出する方法の開発を行ったものである。得られた結果は以下のとおりである。

- (1) ニューラルネットワークを利用した異常検知システムの構築において、ウェーブレットフィルターバンクによるノイズ除去、発酵進行度を示すプロセス変数による規格化、中間ノードの出力のクラスターをもとにした閾値の設定について異常検知システムの性能を向上させる方法を確立している。
- (2) 平均仮説検定を用いて素データセットから識別力のあるプロセス変数を選択し、入力データセットの品質を改善することによりニューラルネットワークなどの多変量解析モデルの情報抽出性能を向上させる方法を確立している。
- (3) 自己組織化マップにより特徴パターンを抽出する際に、統計学的、生物学的見地に基づく評価関数によって適切なニューロンの数を決定する方法を確立している。
- (4) 網羅的遺伝子発現のデータからニューラルネットワークによりクラスター化された遺伝子ネットワークの相互作用を推定する方法を確立している。

このように、本研究は、生物プロセスシステム工学の発展に寄与するところが大きく、よって本論文は、博士論文として価値あるものと認める。