

Częstochowa, dn. 2 czerwca 2016 r.

Dr hab. inż. Rafał Scherer, prof. Politechniki Częstochowskiej
Instytut Inteligentnych Systemów Informatycznych
Wydział Inżynierii Mechanicznej i Informatyki
Politechnika Częstochowska
al. Armii Krajowej 36
42-200 Częstochowa



Recenzja

rozprawy doktorskiej mgr inż. Mateusza Wójcika pt.: Hybrydowy system neuronalny
inteligentnego monitoringu turbiny wiatrowej
Promotor: dr hab. Andrzej Bielecki, prof. AGH

Niniejszą recenzję opracowano na zlecenie Dziekana Wydziału Elektrotechniki, Automatyki,
Informatyki i Inżynierii Biomedycznej z dnia 29.02.2016 r.

1. Charakterystyka tematu, celu i tezy badawczej rozprawy

Diagnostyka maszyn jest niezwykle istotną dziedziną, szczególnie w tak drogich, trudnodostępnych oraz obciążonych urządzeniach, jakimi są turbiny wiatrowe. Farmy wiatrowe stają się coraz powszechniejszym elementem krajobrazu, dlatego też rośnie zapotrzebowanie na systemy potrafiące w sposób predykcyjny dokonywać ich diagnostyki. Pomimo trwających badań światowych, opracowanie metod umożliwiających udowodnienie tezy rozprawy określonej następująco: „przy pomocy narzędzi sztucznej inteligencji możliwe jest stworzenie w pełni autonomicznego systemu, który w czasie rzeczywistym będzie analizować parametry pracy turbiny wiatrowej, wykrywając nietypowe stany operacyjne lub taką zmianę parametrów pomiarowych, która może świadczyć o usterce lub wystąpieniu stanu przedawaryjnego.” jest dalej zadaniem stosunkowo trudnym. Turbiny wyposażone są w dużą liczbę czujników, co powoduje, że dane są wielowymiarowe. Ponadto obciążenie turbiny ma bardzo trudny charakter. Aby zmierzyć się z problemem Autor stworzył nowatorski, rozbudowany system do grupowania wielowymiarowych danych w czasie rzeczywistym oraz opracował rozliczne metody towarzyszące.

2. Zawartość rozprawy

Recenzowana praca magistra inżyniera Mateusza Wójcika składa się z 7 rozdziałów zasadniczych, streszczenia w języku polskim i angielskim, podsumowania, bibliografii oraz dodatku. Dokument liczy 82 strony.

Pierwszy rozdział jest wprowadzeniem do tematyki, zawiera cel, tezę oraz strukturę pracy.

Rozdział 2 jest przybliżeniem metod użytych w pracy. Dokonano taksonomii metod grupowania danych. Podano wymagania stawiane systemom czasu rzeczywistego. Autor opisał bardzo dokładnie sieci ART-2 z architekturą oraz znaczeniem poszczególnych grup wag. Tak dokładny opis jest niespotykany wcześniej w literaturze polskiej. Architektura sieci jest przedstawiona na rysunku 2.2, a nie jak podano w pracy 2.3.1. Rysunek jest tożsamy ze schematem z artykułu Chun-Hsien Chen, Li Pheng Khoo, Wei Yan, "A strategy for acquiring customer requirement patterns using laddering technique and ART2 neural network", *Advanced Engineering Informatics* 16 (3) 2002, ss. 229-240, którego Autor nie cytuje. W Podrozdziale 2.3.6 zostały podane szczegółowe własności sieci ART-2. Podrozdział 2.4 opisuje różne aspekty stosowania sieci ART-2 do grupowania danych w czasie rzeczywistym oraz normalizacji stereograficznej dla polepszenia efektywności klasyfikacji. Autor przedstawił modyfikacje sieci ART-2 oraz dokonał porównania z innymi modyfikacjami spotykanymi w literaturze.

W trzecim rozdziale omówiono klasyfikację binarną. Skupiono się na metodzie Mixture of Gaussians oraz Expectation Maximization. Rozdział ten jest przeglądem literatury. Podano przykłady działania klasyfikatorów obszarowych dla danych z monitoringu turbiny wiatrowej, niestety bez opisu osi, co jednak nie jest wielką wadą w tym rozdziale.

Rozdział 4 pokazuje jak wykorzystać znaną w literaturze metodę Variational Bayesian EM do znalezienia cech szczególnych zbioru danych. Autor dokonał szczegółowej analizy algorytmu oraz przedstawił go w tzw. Plate Notation. Algorytm został wybrany z powodu swojej odporności na osobliwości w zbiorze danych oraz na przeuczenie, co ułatwia użycie tej metody do wyznaczenia optymalnej liczby komponentów w sumie gaussowskiej. Również w tym rozdziale wykresy obrazujące działanie opisywanych metod nie mają opisu osi.

Najważniejsze osiągnięcie, czyli system do grupowania danych w czasie rzeczywistym zostało przedstawione w Rozdziale 5. Wykorzystuje on metody przedstawione w rozdziałach 2-4, a mianowicie sieć neuronową ART-2, metodę projekcji stereograficznej, Mixture of Gaussians oraz Variational Bayes Mixture of Gaussians. System zaprojektowany jest do

przetwarzania ciągłego strumienia danych w czasie rzeczywistym. Zaprezentowany model jest elastyczny i uniwersalny. Może służyć do diagnostyki innych urządzeń, a nie tylko turbin wiatrowych. Na uwagę zasługuje udostępnienie publiczne systemu w serwisie GitHub oraz Matlab File Exchange wraz z instrukcją użytkownika.

Rozdział 6 omawia zagadnienia związane z monitoringiem turbin wiatrowych. W podrozdziale 6.1 podane są podstawowe techniczne informacje związane z turbinami oraz autorski model opisu prędkości wiatru, a w następnym omówione wcześniejsze prace Autora. Na bazie ich doświadczeń powstał system będący przedmiotem rozprawy opisany w poprzednim rozdziale. Należy podkreślić, że problem wiatru jest skomplikowanym i chaotycznym zjawiskiem, przez co należy pochwalić Autora za wprowadzenie modelu wiatru oraz obciążeń. System sam dobiera parametry sieci ART-2 do zmian charakteru danych pomiarowych, co wielce usprawnia proces monitoringu turbiny.

Rozdział 7 jest testem systemu z wykorzystaniem trzech zbiorów danych z rzeczywistej turbiny, w której wystąpiły usterki. Użyto danych wibracyjnych oraz parametrów operacyjnych maszyny. Proponowany system rozpoznał poprawnie różnego rodzaju stany turbiny, w tym również przedawaryjne. Należy również podkreślić, że system działa w pełni autonomicznie, bez potrzeby ustawiania parametrów początkowych. Nie jest jednak jasne jak dokonano podziału na dane treningowe i testowe. Brakuje również oceny działania układu (przykładowo dla klasyfikacji możemy wyznaczyć dokładność, krzywe ROC czy wskaźniki precision i recall). Niestety wykresy prezentowane w tym rozdziale również nie mają opisanych osi.

Ogólnie zasadnicze i oryginalne rezultaty pracy można podsumować następująco:

- wyczerpujący opis oraz nowe rezultaty naukowe dotyczące sieci ART-2,
- rozbudowany system grupowania danych wielowymiarowych w czasie rzeczywistym,
- oryginalne rezultaty związane z diagnostyką turbin wiatrowych, np. model opisu prędkości wiatru.

Wymienione oryginalne metody przedstawione w rozdziałach 2-7 zostały opublikowane w artykułach naukowych w materiałach konferencyjnych oraz czasopismach z tzw. Listy Filadelfijskiej. Zaprezentowany materiał pokazuje, że Doktorant udowodnił tezę pracy.

3. Uwagi krytyczne i wskazówki dotyczące rozprawy

Praca jest napisana dobrze, jednak w czasie czytania nasunęły mi się następujące wątpliwości:

Autor użył w testach w Rozdziale 7 trzech zbiorów danych niedostępnych publicznie. Czy nie było celowe użycie popularnych ogólnodostępnych zbiorów danych w celu porównania z innymi metodami znanymi w literaturze? Przykładem może być zbiór danych Bearing Fault Dataset (dostępny pod adresem <http://www.mfpt.org/FaultData/FaultData.htm>). Bez wątpienia warto było zweryfikować tak rozbudowany i nowatorski system na innych danych.

Czemu system oparty jest właśnie na sieci ART-2, a nie analizie sygnałów czy innych metodach inteligencji obliczeniowej, np. nieliniowe neuronowe sieci wielowarstwowe? Wiąże się to z następnym pytaniem.

Czy można porównać dokładność i wydajność zaproponowanego systemu z innymi znanymi w literaturze?

Oczywiście powyższe uwagi mogą odnosić się do wielu prac naukowych, gdyż ze względu na ograniczony czas każdego przedsięwzięcia naukowego i ogromną liczbę rozwiązań w nauce światowej, nie jest możliwe wyczerpujące przetestowanie konkurencyjnych algorytmów czy zbiorów danych.

W pracy brak jest dokładnego opisu użytych zbiorów danych, poszczególnych zmiennych, oraz sposobu ich użycia, czyli podziału na dane treningowe i testowe.

Patrząc pod kątem edytorskim można zauważyć kilka błędów, np.:

- kilka błędów, tzw. literówek,
- kropka jako separator dziesiętny (str. 60),
- brak opisu osi na większości rysunków.
- brak pochylenia nazw zmiennych w tekście.

Oczywiście celem tak postawionych pytań i uwag nie jest jakiegokolwiek umniejszenie wartości naukowej rozważanej rozprawy doktorskiej.

4. Wnioski końcowe recenzji

Podsumowując recenzję stwierdzam, że Pan mgr inż. Mateusz Wójcik w rozprawie doktorskiej „Hybrydowy system neuronalny inteligentnego monitoringu turbiny wiatrowej”:

- Zrealizował cel i zweryfikował hipotezę rozprawy,
- uzyskał oryginalne rezultaty naukowe dotyczące uczenia maszynowego, a w szczególności sieci ART-2,
- stworzył rozbudowany system grupowania danych wielowymiarowych w czasie rzeczywistym,
- uzyskał oryginalne rezultaty związane z diagnostyką turbin wiatrowych, np. wprowadził model opisu prędkości wiatru,
- wykazał się umiejętnością samodzielnej pracy badawczej, znajomością literatury światowej i bogatą wiedzą w dziedzinie klasyfikacji, grupowania danych oraz sieci ART-2.

Recenzowana praca spełnia wymagania ustawy o tytule i stopniach naukowych. Wnoszę o jej przyjęcie i dopuszczenie do publicznej obrony.

Radosław Skarżewski