

Warszawa, dn. 10 września 2023 roku

mgr inż. Dawid Szpadzik

(stopień, imię i nazwisko osoby ubiegającej się o nadanie stopnia doktora)

Wydział Chemiczny / Katedra Technologii Chemicznej

(Wydział/Katedra/Zakład)

Politechnika Warszawska

(Uczelnia)

STRESZCZENIE ROZPRAWY DOKTORSKIEJ

pt. „Badanie możliwości wykorzystania narzędzi statystycznych w analizach jakościowych w obszarach produkcyjnych”

Promotor: prof. dr hab. inż. Wioletta Raróg-Pilecka

W niniejszej rozprawie przedstawiono tworzenie narzędzia statystycznego wykorzystywanego w analizach jakościowych, w obszarach produkcji chemicznej, z wykorzystaniem jedynie standardowego oprogramowania i komputera biurowego.

Opisano uczenie maszynowe oraz przykłady jego zastosowanie w obszarze technologii chemicznej łącznie z rysem historycznym. Przedstawiono powszechnie stosowane modele klasyfikacji binarnej, takie jak regresja logistyczna, losowy las decyzyjny, drzewa klasyfikacji i inne. Metodę wykorzystaną w niniejszych badaniach (algorytm naiwnego klasyfikatora Bayesa) opisano szerzej oraz zilustrowano ją przykładem. Przytoczono sposoby oceny skuteczności działania klasyfikatorów binarnych wykorzystujące tablicę pomyłek.

W związku z wdrożeniowym charakterem badań opisany został proces produkcji artykułów chemii gospodarczej. Polegał on na wytwarzaniu płynnych środków czystości na bazie roztworów wodnych w modelu szarżowym. Zwrócono uwagę na zarządzanie jakością w zakładzie wytwórczym (normy ISO), której celem było zapewnienie, że wyroby wprowadzane na rynek są zgodne z deklaracjami producenta oraz obowiązującymi normami prawnymi. Uwzględniony został aspekt finansowy (kosztochłonności) utrzymania laboratorium fizykochemicznego, który składał się z wydatków na materiały zużywalne (np. odczynniki, pipety, rękawiczki) oraz zatrudnienie wykwalifikowanego personelu. Praca badawcza została podjęta w ramach działań ciągłego doskonalenia procesu produkcyjnego oraz optymalizacji kosztów wytwarzania w przemyśle chemicznym. Dążyły one ku zapewnieniu odpowiedniej marżowości produktu oraz konkurencyjnej ceny dla konsumenta. Działania te

były szczególnie istotne w kontekście presji inflacyjnej wywołanej pandemią COVID-19 oraz sytuacją geopolityczną w Europie.

W części eksperymentalnej, w pierwszej fazie przeanalizowano proces wytwórczy oraz technologiczny realizowany u współpracującego przedsiębiorcy. Ustalono wymagania stawiane opracowywanemu narzędziu oraz kryteria sukcesu. Wyszczególniono zostało kryterium badawcze (narzędzie klasyfikujące osiąga dokładność lepszą niż losowa) oraz kryterium wdrożeniowe (dokładność przypisania klasy na poziomie minimum 95%). Opisano działanie stworzonej aplikacji oraz wbudowane funkcje, których celem było wypełnienie wymagań postawionych w poprzedzającym etapie. Klasyfikator wyposażono w szereg parametrów, które mogą być modyfikowane przez użytkownika.

Drugim etapem części eksperymentalnej była optymalizacja działania algorytmu uczenia maszynowego (tzw. trenowanie). W pierwszej kolejności został ustalony ogólny punkt odniesienia – ustawienia referencyjne. Następnie każdy parametr algorytmu był iteracyjnie modyfikowany w celu zbadania charakterystyki jego wpływu na jakość klasyfikacji. W kolejnym kroku wysterowano wszystkie argumenty jednocześnie, przy użyciu tych nastaw, dla których uprzednio uzyskano najlepsze wartości dokładności oraz współczynnika korelacji. Ostatnim krokiem było zbadanie możliwości ograniczenia liczby analiz fizykochemicznych, w których zużywane są materiały jednorazowe oraz odczynniki.

Udowodniono, że jest możliwe opracowanie narzędzia wykorzystującego model uczenia maszynowego, pracującego na komputerze osobistym, stworzonego przy użyciu oprogramowania biurowego Microsoft Office. Potwierdzono możliwości wykorzystania narzędzi statystycznych w kontroli jakości w przemyśle chemicznym (cel badawczy pracy doktorskiej). Dokładność opracowanego narzędzia klasyfikującego była większa od 50% – tym samym lepsza od przypisania losowego. Został osiągnięty wymagany do wdrożenia poziom dokładności (minimum 95%), przy jednoczesnym zachowaniu kryterium przyznania ocen próbkom ze zbioru treningowego (minimum 80%). W populacji testów osiągnięta najmniejsza dokładność wyniosła 97,85%. Prace optymalizacyjne wykazały, że liczba kosztochłonnych badań fizykochemicznych może zostać zredukowana, co dodatkowo może wpłynąć na zwiększenie bezpieczeństwa personelu laboratoryjnego.

W toku prac badawczych wykazano, że korekcie matematyczne nie przyniosły pożądanych efektów. Niwelowanie wpływu braku wystąpień obserwacji (wygładzenia Laplace'a) spowodowało znaczny pogorszenie wskaźników jakości klasyfikacji. Nie zaobserwowano również znaczących zmian w metrykach przy zastosowaniu dyskretyzacji zmiennych rzeczywisty w oparciu o odchylenie standardowe.

W przedstawionej rozprawie doktorskiej udowodniono, że jest możliwe zbudowanie algorytmu uczenia maszynowego z wykorzystaniem standardowego oprogramowania (Microsoft Office) oraz komputera biurowego. Aplikacja spełniła wymogi, aby uzyskać potencjał wdrożeniowy w zakładzie wytwórczym przemysłu chemicznego Reckitt Benckiser Production (Poland) sp. z o.o. Dokładność klasyfikacji osiągnęła poziom 99,85%, a współczynnik korelacji 0,9758. Omówiono parametry sterujące algorytmem oraz ich wpływ na skuteczność oznaczenia klasy. Reasumując, uwidoczniło się zastosowanie uczenia maszynowego w technologii produkcji płynnych środków czystości oraz udowodniono, że jego opracowanie nie wymaga specjalistycznych narzędzi komputerowych. Możliwe jest jego wykorzystanie w powszechnych procesach produkcyjnych środków chemii gospodarczej, zmniejszając koszty związane z utrzymaniem laboratorium kontroli jakości. Tym samym optymalizacja działań inżynierów chemików rozszerza się o modelowanie matematyczne całego zakresu wykonywanych zadań w przemyśle chemicznym, nie zaś jedynie na pojedynczych metodach fizykochemicznych.

Dawid Szpadzik

mgr inż. Dawid Szpadzik

Słowa kluczowe: uczenie maszynowe, naiwny klasyfikator Bayesa, klasyfikacja binarna, przemysł chemiczny, kontrola jakości w przemyśle chemicznym, technologia produkcji środków chemii gospodarczej