

Streszczenie pracy doktorskiej mgr Diany Domańskiej pt.: „Prognozowanie eksploracyjne danych wielowymiarowych”

Podstawowym celem rozprawy doktorskiej było opracowanie metody prognozowania eksploracyjnego złożonych danych. Pośrednimi celami było określenie miary podobieństwa dla rzeczywistych, złożonych obiektów, ich klasyfikacja oraz zbudowanie algorytmu zamiany ciągu liczbowego na liczbę rozmytą. Mając na uwadze niewielką ilość danych oraz to, że posiadane dane pochodzą z zaledwie kilku stacji pomiarowych, które są bardzo odległe od siebie zastosowano do rozwiązania problemu podejście oparte na podobieństwie danych. Chaotyczny charakter prognozowanego zjawiska tj. jego niejednoznaczna, ciągła i nieprecyzyjna natura oraz nieprecyzyjność i niekompletność danych sprawiła, że metoda prognozowania oparta jest na teorii zbiorów rozmytych.

W przypadku wielu prognozowanych zjawisk posiadamy pewną wiedzę o zachowaniu się w przeszłości zjawisk, które mają istotny wpływ na zjawisko analizowane. W pierwszej kolejności, chcąc oprzeć prognozę na informacji o zjawiskach współistniejących, należy zbadać zależność pomiędzy zjawiskiem prognozowanym a zjawiskami współistniejącymi. Jeśli zależność taka występuje, to w prognozie badanego zjawiska będzie można wykorzystać prognozę zjawisk współistniejących. Dlatego też jeśli znamy historię zdarzeń zjawisk współistniejących, to możemy w przeszłości odszukać podobne (analogiczne) przebiegi zjawisk współistniejących do ich prognozy. Następnie z podobnych przebiegów zjawisk współistniejących z przeszłości otrzymujemy potencjalne przebiegi prognozowanego zjawiska na przyszłość. W ostatnim etapie prognozy te są agregowane. Taki sposób przewidywania określonego zjawiska w oparciu o wiedzę o pewnych skorelowanych ze sobą zjawiskach pochodząca z przeszłości, teraźniejszości i czasami przyszłości nazywamy *prognozowaniem eksploracyjnym*.

Zaproponowana metoda prognozowania eksploracyjnego przetestowana została w prognozowaniu emisji wybranych zanieczyszczeń powietrza: PM_{10} , SO_2 , CO , NO_2 i O_3 . W badaniach użyta została baza autorska wypełniona: numerycznymi prognozami pogody z modelu COSMO LM, parametrami meteorologicznymi mierzonymi na stacjach synoptycznych, stężeniami zanieczyszczeń zmierzonymi na stacjach monitoringowych, informacjami na temat kierunku wiatru wiejącego z poszczególnych sektorów stacji monitoringowych oraz informacjami na temat podobieństwa stacji monitoringowych.

Dane potrzebne do eksperymentów pochodziły z lat 2005-2012. Dane z okresu 2005-2010 służyły jako dane historyczne, z których w eksperymentach korzystała metoda. A dane z okresu 2011-2012, pochodzące z 4 różnych stacji monitoringowych znajdujących się w Złotym Potoku, Cieszynie, Katowicach i Wodzisławiu Śląskim, służyły do weryfikacji metody. Wyniki otrzymane podczas eksperymentów analizowane były odrębnie dla każdego predykowanego stężenia zanieczyszczenia i oceniane były za pomocą miar: przedziałowej, $RMSE$ i MAE .

Uwzględniając wyniki przeprowadzonych eksperymentów i ich porównanie z metodami z literatury można powiedzieć, że prognozowanie eksploracyjne jest efektywną metodą prognozowania.