

## **Popularnonaukowe streszczenie projektu**

Żyjemy w dobie natłoku informacji. Dane opisujące różne zjawiska i zależności dostępne są na wyciągnięcie ręki. Ich zdobycie i gromadzenie nie stanowi większego problemu wobec rosnącej pojemności systemów bazodanowych i mocy obliczeniowej komputerów. Wyzwaniem natomiast jest wydobycie z tych danych użytecznej wiedzy, która pozwoli szybko zdiagnozować pacjenta, opracować strategię działania firmy, rozpoznawać obraz z kamery czy prognozować pogodę. We wszystkich tych przypadkach eksploracji danych wykorzystuje się narzędzia statystyki i uczenia maszynowego. Sztuczne sieci neuronowe należą do najpopularniejszych metod uczenia maszynowego i są powszechnie stosowane do budowy systemów decyzyjnych, rozpoznawania obrazów, modelowania zjawisk fizycznych, prognozowania i sterowania. Wiedzę pozyskują z danych w procesie uczenia się i zapamiętują w swoich wewnętrznych parametrach - wagach synaptycznych. Jednak z uwagi na warstwową strukturę sieci neuronowych proces uczenia jest skomplikowany, bywa niewydajny i nie zawsze prowadzi do dobrych rezultatów. Algorytmy uczenia sieci neuronowych oparte są na metodach gradientowych, które są czasochłonne, wrażliwe na początkowe wartości parametrów i zbiegają się do minimów lokalnych funkcji celu zamiast globalnych.

W ostatnich latach intensywnie rozwijają się alternatywne metody uczenia sieci neuronowych, tzw. metody randomizowane, w których wagi neuronów w wewnętrznej warstwie sieci (warstwie ukrytej) wybierane są losowo, a wagi w warstwie wyjściowej ustalane są przy użyciu prostego rachunku macierzowego. Takie podejście prowadzi do uproszczenia implementacji i pozwala kilkusetkrotnie skrócić czas uczenia przy zachowaniu dokładności modelowania. Wielu badaczy stosujących randomizowane metody uczenia zwraca uwagę na właściwy sposób generowania losowych wag: gdy wagi losowane są z niewłaściwego zakresu, wyniki modelowania są znacznie gorsze. W literaturze brak jednak wskazówek w jaki sposób prawidłowo generować wagi.

W ramach tego projektu planuje się opracować nowe metody generowania losowych parametrów sieci neuronowych. Na podstawie badań teoretycznych dotyczących aproksymacji funkcji i komponowania krzywej aproksymującej poprzez agregację funkcji aktywacji neuronów, wyprowadzone zostaną wzory na obliczanie parametrów funkcji aktywacji. Funkcje aktywacji neuronów zostaną rozmieszczone w przestrzeni w taki sposób, aby ich najbardziej nieliniowe fragmenty wstawiane były w obszarach wypełnionych przez dane. Pozwoli to skutecznie modelować silnie nieliniowe zależności z dużą dokładnością, przy maksymalnym wykorzystaniu zasobów (przy obecnie używanych randomizowanych metodach uczenia wiele neuronów nie bierze udziału w modelowaniu, ponieważ ich funkcje aktywacji mają nasycenie w rejonach wypełnionych danymi i z tego względu są nieprzydatne do modelowania nieliniowych zależności). Rezultatem badań będą nowe algorytmy generowania losowych parametrów sieci neuronowych dla problemów regresji oraz klasyfikacji. Pozwolą one poprawić dokładność modelowania i generalizację.