

Temat:

Procedura wyznaczania rozkładu zbioru danych o dużej wymiarowości

Opiekun naukowy: prof. dr hab. inż. Piotr Kulczycki  
kulczycki@agh.edu.pl

Dyscyplina: informatyka (w dziedzinie nauk technicznych)

Przedmiotem pracy jest implementacja metod eksploracji danych, zwłaszcza analizy skupień, do stworzenia aplikacyjnej procedury wyznaczenia rozkładu zbioru danych, uniezależnionej od typu tego rozkładu. Rozważany będzie przypadek dominujący we współczesnej praktyce: zbiorów licznych i bardzo licznych oraz strumieniowych, a także o współrzędnych ciągłych i kategoriowych (nominalnych oraz porządkowych) przy dużej wymiarowości, co stanie się głównym przedmiotem badań.

Wyniki zostaną zweryfikowane za pomocą ilustracyjnych danych syntetycznych oraz wzorcowych benchmarków, a także ewentualnie na danych rzeczywistych w tematyce zaproponowanej przez Doktoranta.

Przewód, w dyscyplinie informatyka, w dziedzinie nauk technicznych, prowadzony będzie w Instytucie Badań Systemowych Polskiej Akademii Nauk w Warszawie.

Wskazana jest biegłość programowania oraz predyspozycje matematyczne.