

**Szkolenie pn. „Data mining- kurs podstawowy”
realizowane dla Uniwersytetu Śląskiego w Katowicach**

Abstrakt

Data mining (inaczej uczenie maszynowe, ang. *machine learning*) jest działem sztucznej inteligencji, coraz częściej stosowanym w nauce, badaniach innowacyjnych i rozmaitych zastosowaniach komercyjnych. Dzięki narzędziom *data mining* można wykrywać anomalie, przewidywać, grupować obiekty w podobne segmenty: ogólnie uzyskiwać odpowiedzi na rozmaite pytania. Przykładowe zastosowania to określania, czy pewien obszar lasu jest zakażony szkodnikiem, jaki jest w pływ zanieczyszczeń w górnym odcinku rzeki na środowisko w dolnym, określanie ryzyka kredytowego, przewidywanie zapotrzebowania na energię elektryczną, określanie właściwości produktu na podstawie danych o jego wytwarzaniu, przewidywanie awarii – zakres możliwych zastosowań jest bardzo szeroki.

Celem kursu jest zapoznanie jego uczestników z podstawowymi zasadami data mining (zgłębiania danych) oraz metodami uczenia maszyn (m.in. drzewami decyzyjnymi i sieciami neuronowymi), tak aby mogli samodzielnie realizować projekty data mining.

Szczegółowe cele szkolenia:

- Poznanie podstawowych zasad data mining
- Uzyskanie podstawowej wiedzy w dziedzinie przygotowania danych do analiz
- Zapoznanie się z najczęściej wykorzystywanymi metodami data mining
- Uzyskanie umiejętności wykonywania analiz data mining w środowisku Statistica

Efekty kształcenia i korzyści dla uczestnika:

- Uczestnik rozumie podstawowe zasady data mining
- Uczestnik ma podstawową wiedzę o przygotowaniu danych do analiz
- Uczestnik ma podstawową wiedzę o najpopularniejszych metodach data mining
- Uczestnik umie przeprowadzić analizę data mining w środowisku Statistica

Forma realizacji

Szkolenie będzie zawierało formy wykładu, prezentacji z praktycznymi przykładami, case studies, dyskusji, wymiany doświadczeń między uczestnikami oraz praktycznych ćwiczeń przy komputerach z wykorzystaniem oprogramowania Statistica. Ćwiczenia zostaną przeprowadzone na przykładowych zbiorach danych Wykonawcy

Program szkolenia:

1. Podstawowe idee data mining
 - Modele data mining
 - Rodzaje zadań data mining
 - Przeuczenie i podział na próby
2. Środowisko Statistica Data Miner
 - Wprowadzenie
 - Dane wejściowe
 - Przeglądarka węzłów
 - Przykład prostej analizy w systemie Statistica Data Miner
3. Specjalistyczne moduły systemu Statistica Data Miner (przegląd)
 - Wstępne przetwarzanie danych
 - Metody predykcyjnego data mining (uczenie z nauczycielem)
 - Odkrywanie wiedzy (uczenie bez nauczyciela)
 - Moduły stosowane po uzyskaniu modelu
4. Wprowadzenie do wybranych metod data mining
 - Regresja logistyczna
 - Drzewa klasyfikacyjne i regresyjne
5. Przykłady analiz w systemie Statistica Data Miner
 - Wstępna obróbka danych – czyszczenie i przekształcenia
 - Zadanie regresyjne
 - Analiza skupień (segmentacja)
 - Problem klasyfikacyjny – modelowanie zdolności kredytowej, zadanie regresyjne, analiza skupień (segmentacja).

Informacje o wykonawcy i trenerach realizujących szkolenie

StatSoft Polska Sp. z o.o.
ul. Kraszewskiego 36, 30-110 Kraków
tel. 12 428 43 00, e-mail: info@statsoft.pl
www.StatSoft.pl

Wykładowca: mgr Tomasz Demski

Projekt pt. „Jeden Uniwersytet – Wiele Możliwości. Program Zintegrowany”

Harmonogram szkolenia

Szkolenie/Warsztaty
„Data mining- kurs podstawowy”
realizowane dla Uniwersytet Śląski w Katowicach w terminie:.....

Czas	Moduł, zagadnienia
I DZIEŃ	
Godz. 09:00 – 16:00 (8 godz. dydaktycznych) +60 min. na przerwy dziennie	- Podstawowe idee data mining - Modele data mining - Rodzaje zadań data mining - Przeuczenie i podział na próby - Środowisko Statistica Data Miner - Wprowadzenie - Dane wejściowe - Przeglądarka węzłów - Przykład prostej analizy w systemie Statistica Data Miner - Specjalistyczne moduły systemu Statistica Data Miner (przegląd) - Wstępne przetwarzanie danych - Metody predykcyjnego data mining (uczenie z nauczycielem) - Odkrywanie wiedzy (uczenie bez nauczyciela) - Moduły stosowane po uzyskaniu modelu
II DZIEŃ	
Godz. 09:00 – 16:00 (8 godz. dydaktycznych) +60 min. na przerwy dziennie	- Wprowadzenie do wybranych metod data mining - Regresja logistyczna - Drzewa klasyfikacyjne i regresyjne - Przykłady analiz w systemie Statistica Data Miner - Wstępna obróbka danych – czyszczenie i przekształcenia - Zadanie regresyjne - Analiza skupień (segmentacja) - Problem klasyfikacyjny – modelowanie zdolności kredytowej, zadanie regresyjne, analiza skupień (segmentacja).