

**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](mailto:info@statsoft.pl)  
[www.StatSoft.pl](http://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                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|-----------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>I DZIEŃ</b>                                                                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| Godz. 09:00 – 16:00<br>(8 godz. dydaktycznych)<br>+60 min. na przerwy<br>dziennie | <ol style="list-style-type: none"> <li>Podstawowe idee data mining <ul style="list-style-type: none"> <li>Modele data mining</li> <li>Rodzaje zadań data mining</li> <li>Przeuczenie i podział na próby</li> </ul> </li> <li>Środowisko Statistica Data Miner <ul style="list-style-type: none"> <li>Wprowadzenie</li> <li>Dane wejściowe</li> <li>Przeglądarka węzłów</li> <li>Przykład prostej analizy w systemie Statistica Data Miner</li> </ul> </li> <li>Specjalistyczne moduły systemu Statistica Data Miner (przegląd) <ul style="list-style-type: none"> <li>Wstępne przetwarzanie danych</li> <li>Metody predykcyjnego data mining (uczenie z nauczycielem)</li> <li>Odkrywanie wiedzy (uczenie bez nauczyciela)</li> <li>Moduły stosowane po uzyskaniu modelu</li> </ul> </li> </ol> |
| <b>II DZIEŃ</b>                                                                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| Godz. 09:00 – 16:00<br>(8 godz. dydaktycznych)<br>+60 min. na przerwy<br>dziennie | <ol style="list-style-type: none"> <li>Wprowadzenie do wybranych metod data mining <ul style="list-style-type: none"> <li>Regresja logistyczna</li> <li>Drzewa klasyfikacyjne i regresyjne</li> </ul> </li> <li>Przykłady analiz w systemie Statistica Data Miner <ul style="list-style-type: none"> <li>Wstępna obróbka danych – czyszczenie i przekształcenia</li> <li>Zadanie regresyjne</li> <li>Analiza skupień (segmentacja)</li> <li>Problem klasyfikacyjny – modelowanie zdolności kredytowej, zadanie regresyjne, analiza skupień (segmentacja).</li> </ul> </li> </ol>                                                                                                                                                                                                                |