

## Studia podyplomowe w zakresie przetwarzanie, zarządzania i statystycznej analizy danych

PRZEDMIOT (liczba godzin konwersatoriów/ćwiczeń)	WYMAGANE TREŚCI MERYTORYCZNE ZAJĘĆ
<p>Statystyka opisowa z elementami analizy regresji (4/19)</p>	<ol style="list-style-type: none"> <li>1. Raportowanie wyników analiz</li> <li>2. Graficzna prezentacja danych za pomocą różnego rodzaju wykresów</li> <li>3. Odczytywanie kluczowych informacji o danych na podstawie wykresów</li> <li>4. Wyznaczanie i interpretacja podstawowych miar statystyki opisowej</li> <li>5. Omówienie podstawowych związków pomiędzy miarami statystyki opisowej a graficzną reprezentacją danych</li> <li>6. Korelacja i jej własności ( omówienie teoretycznych aspektów problemu poparte będzie dużą liczbą obrazowych przykładów)</li> <li>7. Omówienie sposobów wyznaczania linii regresji, porównywanie różnych rodzajów linii regresji</li> <li>8. Regresja logistyczna</li> <li>9. Omówienie własności i zasad budowy współczynnika korelacji rang Spearmana</li> <li>10. Omówienie współczynników dopasowania modelu do danych empirycznych</li> </ol> <p>Narzędzia używane podczas zajęć: MS Excel, Statistica, SPSS, pakiet R</p>
<p>Wnioskowanie statystyczne (4/16)</p>	<ol style="list-style-type: none"> <li>1. Model statystyczny             <ol style="list-style-type: none"> <li>1.1. Pojęcie modelu statystycznego</li> <li>1.2. Przykłady wybranych modeli: dwumianowy, Poissona, wykładniczy, gaussowski</li> <li>1.3. Podstawowe twierdzenia statystyki matematycznej</li> </ol> </li> <li>2. Estymacja punktowa             <ol style="list-style-type: none"> <li>2.1. Pojęcie estymatora</li> <li>2.2. Kryteria estymacji (nieobciążoność, wariancja, ryzyko średniokwadratowe)</li> <li>2.3. Techniki estymacji: metoda największej wiarygodności, metoda najmniejszych kwadratów</li> </ol> </li> <li>3. Przedziały ufności             <ol style="list-style-type: none"> <li>3.1. Pojęcie przedziału ufności</li> <li>3.2. Metody konstrukcji przedziałów ufności</li> <li>3.3. Przykłady konstrukcji przedziałów ufności</li> </ol> </li> <li>4. Weryfikacja hipotez statystycznych             <ol style="list-style-type: none"> <li>4.1. Podstawowe pojęcia: hipoteza statystyczna, test statystyczny</li> <li>4.2. Poziom istotności, poziom krytyczny, hipoteza alternatywna, moc testu</li> </ol> </li> <li>5. Zagadnienia wybrane: podejmowanie decyzji statystycznych, elementy statystyki bayesowskiej</li> </ol> <p>Narzędzia używane podczas zajęć: MS Excel, Statistica, SPSS, pakiet R</p>



<p>Analiza szeregów czasowych (2/8)</p>	<ol style="list-style-type: none"> <li>1. Omówienie celu analizy szeregów czasowych</li> <li>2. Omówienie struktury szeregów czasowych (trend, wahanie sezonowe, wahanie cykliczne, wahanie przypadkowe)</li> <li>3. Omówienie podstawowych modeli matematycznych ( model addytywny i model multiplikatywny)</li> <li>4. Szereg czasowy bez trendu i szereg czasowy z trendem</li> <li>5. Analiza sezonowości</li> <li>6. Określenie autokorelacji</li> </ol> <p>Narzędzia używane podczas zajęć: MS Excel, Statistica, SPSS, pakiet R</p>
<p>Wielowymiarowa analiza danych (4/12)</p>	<ol style="list-style-type: none"> <li>1. Wielowymiarowa analiza regresji jako narzędzie opisu zależności przyczynowo-skutkowych w gospodarce. Dobór zmiennych. Dobór postaci modelu.</li> <li>2. Modele wielorównaniowe i analiza mnożnikowa. Konstruowanie scenariuszy rozwoju</li> <li>3. Modele jakościowej zmiennej objaśnianej (liniowy, logitowy, probitowy)</li> <li>4. Wielowymiarowa analiza porównawcza jako narzędzie opisu zjawisk złożonych. Statyczna WAP (ustalenie hierarchii „wielowymiarowych” obiektów). Dynamiczna WAP (analiza zmian zjawisk złożonych w czasie). Klasyfikacja obiektów na podstawie wartości miernika syntetycznego</li> <li>5. Analiza skupień –wyodrębnianie jednorodnych grup obiektów. Metody aglomeracyjne</li> <li>6. Dendryt wrocławski. Wybrane metody podziałowe. Ustalanie liczby skupień</li> <li>7. Wielowymiarowa analiza wariancji i jej zastosowania</li> <li>8. Analiza dyskryminacyjna jako narzędzie klasyfikacji obiektów „wielowymiarowych”.</li> <li>9. Własności i ocena jakości funkcji dyskryminacyjnej. Ocena trafności klasyfikacji</li> <li>10. Funkcje dyskryminacyjne dla więcej niż dwóch populacji</li> <li>11. Wielokryterialne modele jako narzędzie podejmowania decyzji (modele Electra, Sparte itp.)</li> <li>12. Analiza głównych składowych i jej zastosowania (wieloczynnikowy model rynku kapitałowego, delimitacja)</li> <li>13. Analiza korespondencji</li> </ol> <p>Narzędzia używane podczas zajęć: MS Excel, Statistica, SPSS, pakiet R</p>
<p>Wybrane metody prognozowania (4/10)</p>	<ol style="list-style-type: none"> <li>1. Omówienie podstawowych pojęć związanych z prognozowaniem</li> <li>2. Klasyfikacja prognoz</li> <li>3. Ocena wiarygodności prognoz</li> <li>4. Liniowe modele zależne od wielu zmiennych</li> <li>5. Nieliniowe model zależne od jednej i wielu zmiennych</li> <li>6. Modele autoregresyjne i średnia ruchoma</li> <li>7. Przegląd wybranych metod taksonomicznych</li> <li>8. Analiza złożonego przypadku</li> </ol> <p>Narzędzia używane podczas zajęć: MS Excel, Statistica, SPSS, pakiet R</p>



<p>Data mining (4/20)</p>	<ol style="list-style-type: none"> <li>1. Podstawowe idee data mining: <ol style="list-style-type: none"> <li>1.1. Modele data mining</li> <li>1.2. Rodzaje zadań data mining</li> <li>1.3. Przeuczenie i podział na próby</li> </ol> </li> <li>2. Środowisko STATISTICA Data Miner <ol style="list-style-type: none"> <li>2.1. Wprowadzenie</li> <li>2.2. Dane wejściowe: lokalne pliki danych i zdalne źródła danych</li> <li>2.3. Przeglądarka węzłów</li> <li>2.4. Przykład prostej analizy w systemie STATISTICA Data Miner</li> </ol> </li> <li>3. Dane zewnętrzne <ol style="list-style-type: none"> <li>3.1. STATISTICA Query</li> <li>3.2. IDP - In-place Database Processing</li> </ol> </li> <li>4. Specjalistyczne moduły systemu STATISTICA Data Miner (przegląd) <ol style="list-style-type: none"> <li>4.1. Wstępne przetwarzanie danych</li> <li>4.2. Metody predykcyjnego data mining (uczenie z nauczycielem)</li> <li>4.3. Odkrywanie wiedzy (uczenie bez nauczyciela)</li> <li>4.4. Moduły stosowane po uzyskaniu modelu</li> </ol> </li> <li>5. Wstępna obróbka danych - czyszczenie i przekształcenia <ol style="list-style-type: none"> <li>5.1. Ogólne informacje</li> <li>5.2. Czyszczenie danych</li> <li>5.3. Niezgodności danych z rzeczywistością</li> <li>5.4. Postępowanie z brakami danych</li> <li>5.5. Dane nietypowe (odstające) i zaszumione</li> <li>5.6. Przekształcenia danych</li> <li>5.7. Redukcja danych</li> <li>5.8. Zmniejszenie liczby wymiarów</li> <li>5.9. Zamiana zmiennej ciągłej na skategoryzowaną (kategoryzacja)</li> <li>5.10. Łączenie klas zmiennej skategoryzowanej</li> <li>5.11. Losowe próbkowanie</li> <li>5.12. Dane anachroniczne</li> </ol> </li> </ol> <p>Narzędzia używane podczas zajęć: Statistica Data Miner, SPSS, pakiet R</p>
<p>Studium analizy przypadku – projekt (0/22)</p>	<p>w trakcie zajęć studenci będą opracowywać złożony problem wymagający wykorzystania wiedzy zdobytej na wielu wcześniejszych zajęciach. Na podstawie rzeczywistych danych pochodzących od słuchacza, lub zaproponowanych przez prowadzącego uczestnicy zajęć będą musieli przygotować plan analizy statystycznej, następnie opracować dane, przygotować raporty oraz opracować sprawozdanie podsumowujące całość rozważań.</p>



<p>Przygotowanie danych na potrzeby analizy statystycznej (0/10)</p>	<ol style="list-style-type: none"> <li>1. Czyszczenie danych <ol style="list-style-type: none"> <li>1.1. Analityczne i graficzne wykrywanie obserwacji błędnych i nietypowych</li> <li>1.2. Obsługa i zastępowanie braków danych</li> <li>1.3. Identyfikacja i usuwanie powtórzonych rekordów</li> </ol> </li> <li>2. Przekształcenia danych <ol style="list-style-type: none"> <li>2.1. Obliczanie zmiennych pochodnych</li> <li>2.2. Zmiana przypisania kategorii (przekodowania)</li> <li>2.3. Przypisywanie rang</li> <li>2.4. Zmienne opóźnione</li> <li>2.5. Operacje na wartościach tekstowych</li> </ol> </li> <li>3. Wykorzystanie selekcji przypadków i stanów przypadków</li> <li>4. Zmiana układu danych <ol style="list-style-type: none"> <li>4.1. Scalanie plików</li> <li>4.2. Przekształcenia Ułoż w stertę i Rozrzuc po zmiennych (operacje na danych typu LIMS)</li> </ol> </li> <li>5. Losowanie i próbkowanie</li> </ol> <p>Narzędzia używane podczas zajęć: MS Excel, Statistica, SPSS, pakiet R</p>
<p>Systemy informacji przestrzennej GIS (4/8)</p>	<ol style="list-style-type: none"> <li>1. Definicja i koncepcje GIS, zastosowania GIS</li> <li>2. Architektura oprogramowania GIS i jego wykorzystanie</li> <li>3. Korzystanie z Internetu i technologii informatycznych w celu pozyskiwania danych GIS</li> <li>4. Zapis cyfrowy środowiska przyrodniczego, dane wektorowe i rastrowe</li> <li>5. Metody pozyskiwania danych przestrzennych i ich atrybutów</li> <li>6. Techniki przetwarzania danych przestrzennych</li> <li>7. Wizualizacja danych na mapie</li> <li>8. Analiza danych przestrzennych</li> <li>9. Geostatystyka</li> </ol> <p>Narzędzia używane podczas zajęć: zależne od prowadzącego ArcGIS, pakiet R</p>
<p>Wprowadzenie do analizy obrazów (4/16)</p>	<ol style="list-style-type: none"> <li>1. Wprowadzenie do środowisk programów do przetwarzania analizy obrazów</li> <li>2. Podstawowe wiadomości z dziedziny pracy z obrazami cyfrowymi</li> <li>3. Akwizycja obrazów</li> <li>4. Wybrane zagadnienia analizy obrazów medycznych</li> <li>5. Poprawa jakości obrazów</li> <li>6. Obrazy w skali szarości</li> </ol> <p>Narzędzia używane podczas zajęć: pakiet do analizy obrazów, pakiet R, Statistica</p>



Zaawansowane techniki analizy obrazów (4/16)	<ol style="list-style-type: none"> <li>1. Podstawowe techniki detekcji</li> <li>2. Segmentacja</li> <li>3. Filtracja</li> <li>4. Przekształcenia morfologiczne</li> <li>5. Szkieletyzacja</li> <li>6. Pogłębiona analiza przykładowych obrazów</li> <li>7. Rozpoznawanie obrazów na przykładzie automatycznego rozpoznawania twarzy</li> </ol> <p>Narzędzia używane podczas zajęć: pakiet do analizy obrazów, pakiet R, Statistica</p>
Zastosowanie sztucznej inteligencji w analizie danych (4/12)	<ol style="list-style-type: none"> <li>1. Sztuczne sieci neuronowe <ol style="list-style-type: none"> <li>1.1. Model neuronu</li> <li>1.2. Pojęcie i działanie sieci neuronowej</li> <li>1.3. Uczenie sieci neuronowej</li> <li>1.4. Zastosowania sieci neuronowych w analizie danych</li> </ol> </li> <li>2. Zbiory rozmyte <ol style="list-style-type: none"> <li>2.1. Pojęcie zmiennej rozmytej</li> <li>2.2. Operacje na zmiennych rozmytych</li> <li>2.3. Wnioskowanie rozmyte</li> <li>2.4. Zastosowanie zbiorów rozmytych</li> </ol> </li> <li>3. Algorytmy genetyczne <ol style="list-style-type: none"> <li>3.1. Reprezentacja osobnika</li> <li>3.2. Funkcja przystosowania</li> <li>3.3. Operatory genetyczne</li> <li>3.4. Zastosowania algorytmów genetycznych w analizie danych</li> </ol> </li> </ol> <p>Narzędzia używane podczas zajęć: Statistica, Pakiet R</p>
Zarządzania i administrowanie bazami danych (2/6)	<ol style="list-style-type: none"> <li>1. Podstawowe obiekty bazy danych i zarządzanie nimi.</li> <li>2. Administracja użytkownikami: <ol style="list-style-type: none"> <li>a. Tworzenie użytkowników.</li> <li>b. Przywileje systemowe i obiektowe, role.</li> <li>c. Ograniczenia (profile).</li> </ol> </li> <li>3. Monitorowanie pracy użytkowników.</li> <li>4. Eksport i import danych.</li> <li>5. Tworzenie nowej bazy.</li> <li>6. Optymalizacja zapytań.</li> </ol> <p>Narzędzia używane podczas zajęć: zależne od prowadzącego</p>
Zarządzanie zbiorami danych (4/8)	<ol style="list-style-type: none"> <li>1. Sposoby przechowywania danych</li> <li>2. Import/ Export danych</li> <li>3. Wybór danych do analizy</li> <li>4. Przygotowanie danych do analizy</li> </ol> <p>Narzędzia używane podczas zajęć: MS Access + zależne od prowadzącego</p>
Analiza danych w praktycznych zastosowaniach (8/0)	Wykłady



Podstawy przedsiębiorczości dla branży IT (5/0)	<ol style="list-style-type: none"><li>1. Formy organizacyjno-prawne przedsiębiorstwa (działalność gospodarcza, spółka cywilna, spółki prawa handlowego: osobowe i kapitałowe).</li><li>2. Procedury i wymagania związane z zakładaniem działalności gospodarczej.</li><li>3. Uproszczone formy ewidencyjne (karta podatkowa, ryczałt, PKPiR)</li><li>4. Różne formy zatrudnienia pracownika. Podstawowe przepisy Kodeksu Pracy.</li><li>5. Źródła finansowania przedsiębiorstw. Finansowanie z wykorzystaniem środków z dotacji.</li><li>6. Zarządzanie projektami IT (metodyki brytyjskie, podejście agile, SCRUM).</li></ol>
---	--

