

<b>KARTA OPISU MODUŁU KSZTAŁCENIA</b>		
Nazwa modułu/przedmiotu <b>Eksploracja danych</b>		Kod <b>1010515321010510542</b>
Kierunek studiów <b>Informatyka</b>	Profil kształcenia (ogólnoakademicki, praktyczny) <b>ogólnoakademicki</b>	Rok / Semestr <b>1 / 2</b>
Ścieżka obieralności/specjalność <b>Technologie wytwarzania oprogramowania</b>	Przedmiot oferowany w języku: <b>polski</b>	Kurs (obligatoryjny/obieralny) <b>obligatoryjny</b>
Stopień studiów: <b>II stopień</b>	Forma studiów (stacjonarna/niestacjonarna) <b>niestacjonarna</b>	
Godziny Wykłady: <b>16</b> Ćwiczenia: - Laboratoria: <b>16</b> Projekty/seminaria: -		Liczba punktów <b>5</b>
Status przedmiotu w programie studiów (podstawowy, kierunkowy, inny) (ogólnouczelniany, z innego kierunku) <b>kierunkowy z danego kierunku</b>		
Obszar(y) kształcenia i dziedzina(y) nauki i sztuki <b>nauki techniczne</b>		Podział ECTS (liczba i %) <b>5 100%</b>
<b>Odpowiedzialny za przedmiot / wykładowca:</b>		
<p>Krzysztof Krawiec            email: Krzysztof.Krawiec@cs.put.poznan.pl            tel. (0-61) 665-3061, fax: (0-61) 877 1525            Instytut Informatyki            ul. Piotrowo 2, 60-965 Poznań</p>		
<b>Wymagania wstępne w zakresie wiedzy, umiejętności, kompetencji społecznych:</b>		
1	<b>Wiedza:</b>	Student rozpoczynający ten przedmiot powinien posiadać podstawową wiedzę z algebry liniowej (proste operacje na wektorach i macierzach) oraz programowania matematycznego / optymalizacji (formułowanie i zapisywanie prostych problemów programowania z ograniczeniami i bez ograniczeń).
2	<b>Umiejętności:</b>	Powinien także posiadać umiejętność projektowania, implementowania i testowania prostych programów komputerowych (w dowolnym języku) realizujących podstawowe operacje wektorowo-macierzowe.
3	<b>Kompetencje społeczne</b>	Student powinien jednocześnie rozumieć konieczność poszerzania swojej wiedzy i swoich umiejętności. Ponadto w zakresie kompetencji społecznych student musi prezentować takie postawy jak uczciwość, odpowiedzialność, wytrwałość, ciekawość poznawcza, kreatywność, kultura osobista, szacunek dla innych ludzi.
<b>Cel przedmiotu:</b>		
<ol style="list-style-type: none"> <li>Przekazanie studentom podstawowej wiedzy z zakresu analizy i wizualizacji danych wielowymiarowych oraz szeregów czasowych, a konkretniej, wiedzy dotyczącej nadzorowanego i nienadzorowanego uczenia sieci neuronowych, analizy szeregów czasowych z elementami prognozowania, a także analizy i wizualizację danych wielowymiarowych z wykorzystaniem metody składowych głównych oraz metody skalowania wielowymiarowego.</li> <li>Rozwijanie u studentów umiejętności identyfikowania, formułowania i rozwiązywania problemów badawczych związanych z wymienionymi wyżej metodami, niezbędnego przetwarzania i strukturalizowania danych wejściowych na potrzeby tych metod a także umiejętności interpretowania różnorodnych ich wyników.</li> <li>Przedstawienie przykładowych zastosowań prezentowanych metod do rzeczywistych problemów praktycznych.</li> <li>Kształtowanie u studentów umiejętności pracy zespołowej.</li> </ol>		
<b>Efekty kształcenia i odniesienie do kierunkowych efektów kształcenia</b>		
<b>Wiedza:</b>		
<ol style="list-style-type: none"> <li>ma uporządkowaną, podbudowaną teoretycznie wiedzę ogólną w zakresie algorytmów, grafiki i komunikacji człowiek-komputer sztucznej inteligencji i uczenia maszynowego - [K_W4]</li> <li>ma uporządkowaną, podbudowaną teoretycznie wiedzę ogólną w zakresie algorytmów, grafiki i komunikacji człowiek-komputer sztucznej inteligencji i uczenia maszynowego - [K_W5]</li> <li>ma podbudowaną teoretycznie szczegółową wiedzę związaną z wybranymi zagadnieniami z zakresu informatyki, takimi jak: eksploracja i wizualizacja danych wielowymiarowych - [K_W6]</li> <li>ma wiedzę o trendach rozwojowych i najistotniejszych nowych osiągnięciach w informatyce, w szczególności w eksploracji danych, oraz w wybranych pokrewnych dyscyplinach naukowych 4. zna podstawowe metody, techniki i narzędzia stosowane przy rozwiązywaniu złożonych zadań inżynierskich z obszaru eksploracji i wizualizacji danych - [K_W8]</li> </ol>		
<b>Umiejętności:</b>		

<p>1. potrafi pozyskiwać informacje z literatury, baz danych oraz innych źródeł (w języku ojczystym i angielskim), integrować je, dokonywać ich interpretacji i krytycznej oceny, wyciągać wnioski oraz formułować i wyczerpująco uzasadniać opinie - [K_U1]</p> <p>2. potrafi określić kierunki dalszego uczenia się i zrealizować proces samokształcenia - [K_U5]</p> <p>3. potrafi planować i przeprowadzać eksperymenty, w tym pomiary i symulacje komputerowe, interpretować uzyskane wyniki i wyciągać wnioski - [K_U8]</p> <p>4. potrafi wykorzystać do formułowania i rozwiązywania zadań inżynierskich i prostych problemów badawczych metody analityczne, symulacyjne oraz eksperymentalne - [K_U9]</p> <p>5. potrafi - przy formułowaniu i rozwiązywaniu zadań inżynierskich - integrować wiedzę z różnych obszarów informatyki (a w razie potrzeby także wiedzę z innych dyscyplin naukowych) oraz zastosować podejście systemowe, uwzględniające także aspekty pozatechniczne - [K_U10]</p> <p>6. potrafi formułować i testować hipotezy związane z problemami inżynierskimi i prostymi problemami badawczymi - [K_U12]</p> <p>7. potrafi ocenić przydatność i możliwość wykorzystania nowych osiągnięć (metod i narzędzi) oraz nowych produktów informatycznych - [K_U13]</p> <p>8. potrafi - stosując m.in. koncepcyjnie nowe metody - rozwiązywać złożone zadania informatyczne, w tym zadania nietypowe oraz zadania zawierające komponent badawczy - [K_U25]</p>
<p><b>Kompetencje społeczne:</b></p> <p>1. rozumie, że w informatyce wiedza i umiejętności bardzo szybko stają się przestarzałe - [K_K1]</p> <p>2. zna możliwości dalszego dokształcania się (studia II i III stopnia, studia podyplomowe, kursy i egzaminy przeprowadzane przez uczelnie, firmy i organizacje zawodowe) - [K_K3]</p> <p>3. zna przykłady i rozumie przyczyny wadliwie działających systemów informatycznych, które doprowadziły do poważnych strat finansowych, społecznych lub też do poważnej utraty zdrowia, a nawet życia - [K_K4]</p> <p>4. potrafi odpowiednio określić priorytety służące realizacji określonego przez siebie lub innych zadania - [K_K6]</p>

### Sposoby sprawdzenia efektów kształcenia

Efekty kształcenia przedstawione wyżej weryfikowane są w następujący sposób:

Ocena formująca:

- a) w zakresie wykładów:
- na podstawie odpowiedzi na pytania dotyczące materiału omówionego na poprzednich wykładach
- b) w zakresie laboratoriów / ćwiczeń:
- na podstawie oceny bieżącego postępu realizacji zadań

Ocena podsumowująca:

- a) w zakresie wykładów weryfikowanie założonych efektów kształcenia realizowane jest przez:
- ocenę wiedzy i umiejętności wykazanych na egzaminie pisemnym w formie testu wyboru obejmującego ok. 25 pytań o charakterze teoretyczno-praktycznym; czas przewidziany na napisanie egzaminu: 60 minut; limit punktów: ok. 25 pkt., na uzyskanie oceny pozytywnej trzeba zdobyć minimalnie więcej niż połowę punktów (czyli np. przy limicie równym dokładnie 25 pkt. wartość ta wynosi 13 pkt.).
  - omówienie wyników egzaminu
- b) w zakresie laboratoriów / ćwiczeń weryfikowanie założonych efektów kształcenia realizowane jest przez:
- ocenę wiedzy i umiejętności poprzez przeprowadzenie obejmującego omówione zagadnienia pisemnego sprawdzianu wiedzy i umiejętności, obejmującego zarówno aspekty teoretyczne (pytania problemowe oraz niewielkie zadania) jak i praktyczne (implementacja komputerowa fragmentów rozwiązań i algorytmów); przykładowy sprawdzian może zawierać 10 pytań/zadań i trwać 45 minut, na uzyskanie oceny pozytywnej trzeba zdobyć minimalnie więcej niż połowę punktów (jak wyżej)

### Treści programowe

Przedmiot przedstawia elementy analizy i wizualizacji danych wielowymiarowych, w tym szeregów czasowych. Omawiane zagadnienia obejmują podstawowe pojęcia uczenia maszynowego, nadzorowane i nienadzorowane uczenie sieci neuronowych, analizę szeregów czasowych z elementami prognozowania, a także analizę i wizualizację danych wielowymiarowych

z wykorzystaniem metody składowych głównych oraz metody skalowania wielowymiarowego.

Wykłady.

Podstawowe pojęcia z dziedziny nadzorowanego uczenia maszynowego oraz tworzenia i testowania klasyfikatorów (na przykładzie drzew decyzyjnych). Następnie, w ramach uczenia sieci neuronowych przedstawione zostają następujące zagadnienia: pojęcia wstępne, idea uczenia sieci, nadzorowane uczenie pojedynczego neuronu na przykładzie reguły delta, uczenie sieci warstwowych, algorytm wstecznej propagacji błędów i jego parametry, interpretacja graficzna procesu uczenia, dobór prędkości uczenia i konfiguracji sieci (liczby i wielkości warstw ukrytych), warunki stopu; nienadzorowane uczenie sieci przez konkurencję, algorytm adaptacyjnego kwantowania wektorowego (LVQ), sieć odwzorowania cech istotnych (SOM); przykłady zastosowań sieci neuronowych. Dalej, w ramach analizy szeregów czasowych przedstawione zostają następujące zagadnienia: pojęcia wstępne, przetwarzanie początkowe, dekompozycja szeregu czasowego na składowe, w tym estymacja trendu

i składowych okresowych, estymacja błędów modelu szeregu czasowego, idea prognozowania oraz wykorzystanie sztucznych sieci neuronowych w prognozowaniu. I wreszcie, w ramach analizy i wizualizacji danych wielowymiarowych przedstawione zostają następujące zagadnienia: podstawy rachunku macierzowego, skalowanie wielowymiarowe (metoda reprezentacji danych wielowymiarowych na płaszczyźnie - tworzenie map danych), analiza korespondencji (metoda obrazowania różnic i podobieństw pomiędzy obiektami), elementy analizy składowych głównych i analizy czynnikowej (metody służące do przetwarzania danych wielowymiarowych), a także metody analizy skupień i hierarchizacji obiektów.

Laboratoria.

Zajęcia laboratoryjne odbywają się w sali komputerowej, i są poprzedzone sesją instruktazową na początku semestru.

Program zajęć laboratoryjnych jest podzbiorem zagadnień realizowanych na wykładach. Część z realizowanych ćwiczeń i zadań wykonywana jest indywidualnie, a część w zespołach 2-osobowych. Do ćwiczeń wykorzystywane

są popularne w dziedzinie uczenia maszynowego zbiory testowe (UCI Repository). Początkowe ćwiczenia realizowane są z wykorzystaniem znanego systemu uczenia maszynowego (Weka). Dalsze, bardziej szczegółowe zadania wymagają implementowania pewnych algorytmów (lub ich wybranych fragmentów) w wybranym języku programowania (np. w języku Java); przykładowym zadaniem tej postaci jest samodzielne zaprojektowanie i przetestowanie programu implementującego warstwową sieć neuronową uczonej z użyciem algorytmu wstecznej propagacji błędów.

Część wymienionych wyżej treści programowych realizowana jest w ramach pracy własnej studenta.

Metody dydaktyczne:

- wykład: prezentacja multimedialna, prezentacja ilustrowana przykładami i zadaniami tablicowymi, pokaz multimedialny
- ćwiczenia laboratoryjne: rozwiązywanie zadań i implementowanie rozwiązań, wykonywanie eksperymentów, prezentacje rozwiązań, oprogramowania i wyników, pytania i dyskusja.

#### Literatura podstawowa:

1. Tadeusiewicz R.: Sieci neuronowe. Akademicka Oficyna Wydawnicza, Warszawa 1993.
2. Osowski, S.: Sieci neuronowe w ujęciu algorytmicznym. WNT, Warszawa 1997.
3. Żurada J., Barski M., Jędruch W.: Sztuczne sieci neuronowe. Podstawy teorii i zastosowania. PWN, Warszawa 1996.
4. Box, G.E.P., Jenkins, G.M.: Analiza szeregów czasowych. PWN, Warszawa, 1983.

#### Literatura uzupełniająca:

1. Gately, E.: Sieci neuronowe. Prognozowanie finansowe i projektowanie systemów transakcyjnych. WIG-Press Warszawa, 1999.
2. Siedlecka, U.: Prognozowanie ostrzegawcze w gospodarce. PWE, Warszawa, 1996.
3. Kassyk-Rokicka, H.: Statystyka nie jest trudna. Tom I: Mierniki statystyczne. PWE, Warszawa, 1998.
4. Zeliaś A., Pawełek B., Wanat S.: Prognozowanie ekonomiczne. Teoria, przykłady, zadania. PWN, Warszawa 2004.

### Bilans nakładu pracy przeciętnego studenta

Czynność	Czas (godz.)
----------	--------------

1. udział w zajęciach laboratoryjnych / ćwiczeniach	16
2. przygotowanie do ćwiczeń laboratoryjnych:	10
3. dokończenie (w ramach pracy własnej) sprawozdań z zajęć laboratoryjnych	16
4. udział w konsultacjach związanych z realizacją procesu kształcenia, dotyczących w szczególności: zajęć laboratoryjnych / projektu (mogą być realizowane drogą elektroniczną)	4 20
5. projektowanie, tworzenie, uruchamianie i testowanie oprogramowania (poza zajęciami laboratoryjnymi)	16 20
6. udział w wykładach	20
7. zapoznanie się ze wskazaną literaturą / materiałami dydaktycznymi (10 stron tekstu naukowego = 1 godz.)	2
8. przygotowanie do egzaminu	1
9. obecność na egzaminie	
10. omówienie wyników egzaminu	
<b>Obciążenie pracą studenta</b>	
<b>forma aktywności</b>	<b>godzin</b>
<b>ECTS</b>	
Łączny nakład pracy	125
Zajęcia wymagające bezpośredniego kontaktu z nauczycielem	37
Zajęcia o charakterze praktycznym	62