

POLITECHNIKA KRAKOWSKA IM. TADEUSZA KOŚCIUSZKI

KARTA PRZEDMIOTU

obowiązuje studentów rozpoczynających studia w roku akademickim 2021/2022

Wydział Informatyki i Telekomunikacji

Kierunek studiów: Informatyka

Profil: Ogólnoakademicki

Forma studiów: niestacjonarne

Kod kierunku: I

Stopień studiów: II

Specjalności: Informatyka stosowana

1 INFORMACJE O PRZEDMIOCIE

NAZWA PRZEDMIOTU	Masowe przetwarzanie danych
NAZWA PRZEDMIOTU W JĘZYKU ANGIELSKIM	Massive processing of data
KOD PRZEDMIOTU	WiT I oJIN D11 21/22
KATEGORIA PRZEDMIOTU	Przedmioty specjalnościowe
LICZBA PUNKTÓW ECTS	3.00
SEMESTRY	3

2 RODZAJ ZAJĘĆ, LICZBA GODZIN W PLANIE STUDIÓW

SEMESTR	WYKŁAD	ĆWICZENIA	LABORATORIUM	LABORATORIUM KOMPUTERO- WE	SEMINARIUM	PROJEKT
3	9	0	18	0	0	0

3 CELE PRZEDMIOTU

Cel 1 Zapoznanie studentów z pojęciami, ważniejszymi algorytmami, gotowymi bibliotekami oraz metodami stosowanymi obecnie do przechowywania, przetwarzania, analizy, eksploracji i wizualizacji olbrzymich ilości danych.

Cel 2 Nabycie umiejętności zastosowania dostępnych gotowych bibliotek (tzw. pakietów) stosowanych do przechowywania, przetwarzania, analizy, eksploracji i wizualizacji olbrzymich ilości danych (zwłaszcza za pomocą

internetowych notebooków i dashboardów udostępnionych przez RStudio dzięki bibliotekom rmarkdown i shiny), zarówno na komputerze, jak i online na istniejących darmowych kontach w Chmurze. Dodatkowo, studenci będą posiadali umiejętność poprawnej interpretacji wyników uzyskanych z algorytmów.

Cel 3 CEL społeczny: Celem będzie nabycie umiejętności samodzielnego poszerzania swojej wiedzy i doskonalenia umiejętności w dziedzinie masowego przetwarzania danych, także docenienia wiedzy z tego zakresu w kształtowaniu współczesnej informatyki. Celem będzie też nabycie umiejętności, aby w sposób staranny i terminowy realizować powierzone sobie zadania, oraz aby być gotowym do rozwiązywania problemów ze wspomnianego zakresu, zarówno w ramach pracy indywidualnej jak i grupowej, a także nabycie umiejętności poszukiwania niezbędnej w tym zakresie wiedzy oraz umiejętności pracy w małych zespołach.

4 WYMAGANIA WSTĘPNE W ZAKRESIE WIEDZY, UMIEJĘTNOŚCI I INNYCH KOMPETENCJI

- 1 Podstawowa wiedza z Algebry liniowej
- 2 Podstawowa wiedza ze Statystyki matematycznej
- 3 Podstawowa wiedza z Rachunku prawdopodobieństwa
- 4 Podstawowa znajomość języka angielskiego

5 EFEKTY KSZTAŁCENIA

EK1 Wiedza Student będzie potrafił wytłumaczyć pojęcia oraz stosowane obecnie metody służące do przechowywania, przetwarzania, analizy, eksploracji i wizualizacji olbrzymich ilości danych. Student będzie potrafił wytłumaczyć działanie oraz zinterpretować wyniki ważniejszych algorytmów stosowanych przy przetwarzaniu, analizie, eksploracji i wizualizacji dużych zbiorów danych.

EK2 Umiejętności Student będzie posiadał umiejętność zastosowania do przechowywania, przetwarzania, analizy, eksploracji i wizualizacji olbrzymich ilości danych ważniejszych algorytmów zawartych w wybranych bibliotekach (pakietach) języka R, środowiska zintegrowanego RStudio oraz w ekosystemie Apache SPARK (czyli w tzw. uniwersalnym silniku Big Data), zarówno na komputerze, jak i online na istniejących darmowych kontach w Chmurze.

EK3 Umiejętności Umiejętności: Student będzie potrafił poprzez RStudio łączyć się internetowo zarówno ze SPARKiem oraz z darmowymi serwerami RStudio, wspomagającymi przetwarzanie dużych zbiorów danych w Chmurze. Student będzie również posiadał umiejętność samodzielnego programowania w języku R w środowisku RStudio.

EK4 Kompetencje społeczne Student nabeędzie umiejętności takie, jak: pogłębianie wiedzy w oparciu o doświadczenia praktyczne i literaturę, inspiracja i organizacja zespołów studenckich, współpraca wewnątrz grup, gotowość do abstrakcyjnego myślenia, gotowość do autonomicznego i odpowiedzialnego wykonywania powierzonych zadań, i sprawność komunikowania się. Student nabeędzie także umiejętność samodzielnego poszerzania swojej wiedzy i doskonalenia umiejętności w dziedzinie masowego przetwarzania danych, także docenienia wiedzy z tego zakresu w kształtowaniu współczesnej informatyki, nabeędzie także umiejętność, aby w sposób staranny i terminowy realizować powierzone sobie zadania, oraz do bycia gotowym do rozwiązywania problemów ze wspomnianego zakresu, zarówno w ramach pracy indywidualnej jak i grupowej, a także nabeędzie umiejętność poszukiwania niezbędnej w tym zakresie wiedzy oraz umiejętność pracy w małych zespołach. Student zauważy też potrzebę samokształcenia i potrzebę ciągłego uaktualniania swej wiedzy.

6 TREŚCI PROGRAMOWE

LABORATORIUM		
LP	TEMATYKA ZAJĘĆ OPIS SZCZEGÓŁOWY BLOKÓW TEMATYCZNYCH	LICZBA GODZIN
L1	Instalacja RStudio oraz R oraz ekosystemu Apache SPARK oraz podstawowych pakietów RStudio oraz R, takich jak sparklyr. Proste eksperymenty w środowisku RStudio z wykorzystaniem języka R i próby łączenia się z ekosystemem SPARK.	1
L2	Instalacja bogatej rodziny tidyverse złożonej z kilkunastu nowoczesnych pakietów RStudio oraz eksperymentowanie z tymi pakietami wzorując się na przykładach i tutorialach dostępnych online na stronach: http://r4ds.had.co.nz/ ; https://github.com/tidyverse ; http://tidyverse.org ; https://github.com/hadley/r4ds ; https://github.com/hadley/dplyr ; https://github.com/tidyverse/ggplot2	2
L3	Instalacja pakietu sparklyr, dplyr i magrittr, data.table, DBI i knitr oraz eksperymentowanie z bazodanowymi funkcjami tych pakietów: select(), filter(), arrange(), rename(), mutate(), group_by(), oraz z operatorem pipeline, tj. %>%.	1
L4	Instalacja i eksperymentowanie z pakietami języka R - pozwalającymi na tworzenie internetowych notebooków. Instalowane pakiety - to: magrittr, Rmarkdown i kilka innych mniejszych.	1
L5	Instalacja i eksperymentowanie z pakietami języka R - pozwalającymi na tworzenie internetowych dashboardów. Instalowany pakiet, to pakiet shiny. Eksperymenty będą bazowały na licznych przykładach zawartych w online tutorialach dostępnych na stronach: http://www.rstudio.com/shiny/ ; http://rstudio.github.io/shiny/tutorial/ ; http://www.rstudio.com/shiny/lessons/Intro/	2
L6	Instalacja i eksperymentowanie z pakietami języka R - pozwalającymi na dokonywanie wzajemnych porównań (podobieństwa i odróżnialności) zbiorów tekstowych zawierających olbrzymie ilości danych. Instalowane pakiety - to: tm, lda, topicmodels, RTextTools, wordcloud, RTextTools, tau i kilka innych dodatkowych.	1
L7	Instalacja i eksperymentowanie z pakietami języka R - pozwalającymi na tworzenie grafów i sieci społecznościowych, bazując na zależnościach istniejących w dużych zbiorach danych. Instalowane pakiety - to: graphframe, igraph, rgl, snowfall, network, tmap, sna, SocialNetworks, tmap, spnet i kilka innych dodatkowych.	1
L8	Instalacja i eksperymentowanie z pakietami języka R - pozwalającymi na tworzenie rekomenderów: recommenderlab, arules, readr, plyr.	1
L9	Założenie darmowego konta w Chmurze (oferowanego przez twórców Apache Spark -Databricks) lub przez środowisko programistyczne RStudio. Przegląd tutorialów z przykładowymi eksperymentami w chmurze i ich naśladowanie. Importowanie Notebooków w Chmurze (z adresu podanego w Literaturze) i ich wykonywanie.	1

LABORATORIUM		
LP	TEMATYKA ZAJĘĆ OPIS SZCZEGÓŁOWY BLOKÓW TEMATYCZNYCH	LICZBA GODZIN
L10	Instalacja i eksperymentowanie z pakietem ISLR. Eksperymenty będą bazowały na licznych przykładach zawartych w licznych online lablatoryjnych przykładach i tutorialach dostępnych na stronie, gdzie są wszystkie kody źródłowe z książki [5]: http://www-bcf.usc.edu/gareth/ISL/code.html ; Książka dostępna jest pod adresem: http://www-bcf.usc.edu/gareth/ISL Dodatkowo dużo przykładów autora książki jest na jego stronie: http://web.stanford.edu/hastie/StatLearnSparsity/	2
L11	Eksperymenty w RStudio z wykorzystaniem platformy H2O i ważniejszych pakietów związanych z tą platformą, tj. pakietem h2o, rsparkling, data.table, dplyr, itp.	1
L12	Eksploracja danych strumieniowych (mining data streams); eksperymentowanie z pakietem stream w języku R, do przetwarzania strumieniowego dużych zbiorów danych.	2
L13	Eksperymentowanie z dodatkowymi, wybranymi pakietami: RAdwords (Web Advertising), manipulate, dpmr, itp. Skorzystanie z alternatywnej do pakietu "sparklyr" (z RStudio) możliwości połączenia R'a ze Sparkiem - tj. pakietu SparkR, jako "frontend'u" prowadzącego z R'a do SPARK'a. Eksperymentowanie z operacjami na ramkach danych (tj. DataFrames), stosowanie operacji: select, filter, aggregate, korzystanie z operacji na kolumnach, funkcji collect() i subset().	2

WYKŁAD		
LP	TEMATYKA ZAJĘĆ OPIS SZCZEGÓŁOWY BLOKÓW TEMATYCZNYCH	LICZBA GODZIN
W1	Charakterystyka pojęcia "Big data" i sumaryczne omówienie zagadnień związanych z tym pojęciem. Wprowadzenie do paradygmatu MapReduce. Przedstawienie i opisanie architektury i funkcji, jakie spełnia ekosystem Apache SPARK i jego podsystemy. Wprowadzenie do środowiska RStudio. Opisanie pakietu sparklyr (autorstwa twórców środowiska RStudio), pozwalającego na połączenie się z programem w języku R ze SPARK'iem oraz pozwalającego na wykonywanie dowolnego pakietu R'a w środowisku skalowalnym i rozproszonym. Omówienie pojęcia ramki danych (tj. DataFrame) oraz ważniejszych operacji na dużych zbiorach danych. Przedstawienie i opisanie przykładów skryptów napisanych z wykorzystaniem pakietu sparklyr.	0.5
W2	Opisanie środowiska RStudio, jego ważniejszych pakietów, ich funkcji i operacji, oraz zalet i metody pracy w tym środowisku. Omówienie metod tworzenia notebook'ów - korzystając z pakietów RStudio: magrittr, markdown, R markdown oraz dplyr.	1
W3	Omówienie metod tworzenia dashboard'ów - korzystając z pakietów RStudio: magrittr, markdown, R markdown oraz shiny.	1

WYKŁAD		
LP	TEMATYKA ZAJĘĆ OPIS SZCZEGÓŁOWY BLOKÓW TEMATYCZNYCH	LICZBA GODZIN
W4	Omówienie ważniejszych operacji bazodanowych udostępnionych przez funkcje ważniejszych pakietów bazo-danowych środowiska RStudio, np. operacji select, filter, aggregate, operacji na kolumnach, funkcji collect() oraz subset().	0.5
W5	Opisanie ważniejszych pakietów R'a służących do eksploracji danych tekstowych (text mining): tm, lda, topicmodels, RTextTools, tau, wordcloud. Przedstawienie i opisanie przykładów skryptów napisanych z wykorzystaniem tych pakietów.	1
W6	Omówienie zasady działania systemów rekomendacji. Opisanie ważniejszych pakietów R'a służących jako rekomendery. Przedstawienie i opisanie przykładów skryptów napisanych z wykorzystaniem pakietów: recommenderlab, arules, readr, plyr.	0.5
W7	Opisanie ważniejszych pakietów R'a służących do eksploracji sieci społecznościowych (social networks). Przedstawienie i opisanie przykładów skryptów napisanych z wykorzystaniem pakietów: sna, network, igraph, SocialNetworks, tmap, spnet.	1
W8	Opisanie ważniejszych pakietów R'a służących do eksploracji danych strumieniowych (mining data streams). Przedstawienie i opisanie przykładów skryptów napisanych z wykorzystaniem pakietu stream, do przetwarzania strumieniowego dużych zbiorów danych.	1
W9	Opisanie ważniejszego pakietu R'a służącego do tworzenia webowych ogłoszeń (Web Advertising) - pakiet RAdwords. Przedstawienie i opisanie przykładów skryptów napisanych z wykorzystaniem tego pakietu, oraz omówienie pakietów: manipulate oraz dpmr.	0.5
W10	Opisanie platformy H2O i ważniejszych pakietów związanych z tą platformą, tj. pakietem h2o, rsparkling, data.table, dplyr, itp. Przedstawienie i opisanie przykładów skryptów napisanych z wykorzystaniem tych pakietów.	0.5
W11	Omówienie zasad i reguł potrzebnych do eksperymentowania w Chmurze na dużych zbiorach danych. Omówienie metody tworzenia i importowania notebook'ów w Chmurze.	1
W12	Omówienie pakietu SparkR - jako alternatywnej opcji dawanej przez pakiet "sparklyr" (z RStudio) i umożliwiającej połączenie R'a ze Sparkiem i pozwalającej na wykonywanie obliczeń w języku R w Chmurze w ekosystemie SPARK. Przedstawienie i opisanie przykładów skryptów napisanych z wykorzystaniem pakietu SparkR.	0.5

7 NARZĘDZIA DYDAKTYCZNE

N1 Wykłady

N2 Ćwiczenia laboratoryjne

N3 Prezentacje multimedialne

N4 Konsultacje

N5 Dyskusja

N6 Praca w 2-3 osobowych grupkach

N7 Kanały komunikacji zdalnej: w przypadku realizacji zajęć w trybie zdalnym - wykorzystane zostaną stosowne narzędzia teleinformatyczne

8 OBCIĄŻENIE PRACĄ STUDENTA

FORMA AKTYWNOŚCI	ŚREDNIA LICZBA GODZIN NA ZREALIZOWANIE AKTYWNOŚCI
Godziny kontaktowe z nauczycielem akademickim, w tym:	
Godziny wynikające z planu studiów	27
Konsultacje przedmiotowe	5
Egzaminy i zaliczenia w sesji	0
Godziny bez udziału nauczyciela akademickiego wynikające z nakładu pracy studenta, w tym:	
Przygotowanie się do zajęć, w tym studiowanie zalecanej literatury	15
Opracowanie wyników	8
Przygotowanie raportu, projektu, prezentacji, dyskusji	35
SUMARYCZNA LICZBA GODZIN DLA PRZEDMIOTU WYNIKAJĄCA Z CAŁEGO NAKŁADU PRACY STUDENTA	90
SUMARYCZNA LICZBA PUNKTÓW ECTS DLA PRZEDMIOTU	3.00

9 SPOSOBY OCENY

OCENA FORMUJĄCA

F1 Ćwiczenia praktyczne podczas zajęć laboratoryjnych

F2 Odpowiedzi ustne podczas zajęć laboratoryjnych

F3 Sprawozdania z umiejętności wykorzystania wybranych bibliotek języka R wywoływanych w środowisku RStudio i Apache SPARK

OCENA PODSUMOWUJĄCA

P1 Średnia ważona ocen formujących

WARUNKI ZALICZENIA PRZEDMIOTU

W1 Warunkiem otrzymania zaliczenia z przedmiotu "Masowe przetwarzanie danych" jest zaliczenie każdej formy zajęć oraz uzyskanie pozytywnej oceny z każdego efektu kształcenia. Dodatkowo wyznaczane będą punkty za obecność na zajęciach i aktywne wykonywanie ćwiczeń praktycznych podczas zajęć, za odpowiedzi ustne podczas zajęć laboratoryjnych oraz za oddanie wszystkich zleconych do napisania sprawozdań wykonanych w środowisku programistycznym RStudio i Apache SPARK, z wykorzystaniem wspomnianych bibliotek.

OCENA AKTYWNOŚCI BEZ UDZIAŁU NAUCZYCIELA

B1 Oddanie wszystkich zleconych do napisania sprawozdań wykonanych w środowisku programistycznym RStudio i Apache SPARK. Wszystkie sprawozdania będą punktowane.

KRYTERIA OCENY

EFEKT KSZTAŁCENIA 1	
NA OCENĘ 2.0	Nie opanowanie wymaganych zagadnień nawet w stopniu 50%.
NA OCENĘ 3.0	Opanowanie zagadnienia w stopniu powyżej 50%.
NA OCENĘ 3.5	Opanowanie zagadnienia w stopniu powyżej 60%.
NA OCENĘ 4.0	Opanowanie zagadnienia w stopniu powyżej 70%
NA OCENĘ 4.5	Opanowanie zagadnienia w stopniu powyżej 80%
NA OCENĘ 5.0	Opanowanie zagadnienia w stopniu powyżej 90%
EFEKT KSZTAŁCENIA 2	
NA OCENĘ 2.0	Nie opanowanie wymaganych zagadnień nawet w stopniu 50%.
NA OCENĘ 3.0	Opanowanie zagadnienia w stopniu powyżej 50%.
NA OCENĘ 3.5	Opanowanie zagadnienia w stopniu powyżej 60%.
NA OCENĘ 4.0	Opanowanie zagadnienia w stopniu powyżej 70%.
NA OCENĘ 4.5	Opanowanie zagadnienia w stopniu powyżej 80%.
NA OCENĘ 5.0	Opanowanie zagadnienia w stopniu powyżej 90%.
EFEKT KSZTAŁCENIA 3	
NA OCENĘ 2.0	Nie opanowanie wymaganych zagadnień nawet w stopniu 50%.
NA OCENĘ 3.0	Opanowanie zagadnienia w stopniu powyżej 50%.
NA OCENĘ 3.5	Opanowanie zagadnienia w stopniu powyżej 60%.
NA OCENĘ 4.0	Opanowanie zagadnienia w stopniu powyżej 70%.
NA OCENĘ 4.5	Opanowanie zagadnienia w stopniu powyżej 80%.
NA OCENĘ 5.0	Opanowanie zagadnienia w stopniu powyżej 90%.
EFEKT KSZTAŁCENIA 4	
NA OCENĘ 2.0	Student nie spełnia warunków określonych dla oceny 3.0

NA OCENĘ 3.0	Student wykonuje powierzone zadania indywidualnie - rezygnuje ze współpracy grupowej i fachowych źródeł wiedzy kosztem jakości rozwiązania. Prace studenta cechuje dopuszczalna niedbałość.
NA OCENĘ 3.5	Student wykonuje powierzone zadania indywidualnie - rezygnuje ze współpracy grupowej i fachowych źródeł wiedzy kosztem jakości rozwiązania. Jego prace są wykonane w sposób staranny.
NA OCENĘ 4.0	Student wykonuje powierzone zadania indywidualnie, a także stara się nawiązać współpracę grupową i sięga po fachowe źródła wiedzy. Współpraca grupowa oraz fachowa literatura mają ograniczony wpływ na wypracowane rozwiązania. Prace studenta zawierają drobne błędy.
NA OCENĘ 4.5	Student wykonuje powierzone zadania indywidualnie, a także stara się nawiązać współpracę grupową i sięga po fachowe źródła wiedzy. Student dostrzega zyski płynące ze współpracy grupowej, jednak jego zbyt małe zaangażowanie powoduje drobne błędy w realizacji projektów.
NA OCENĘ 5.0	Student potrafi rozwiązywać problemy indywidualnie jak i grupowo; korzysta z fachowych źródeł wiedzy; dostrzega zyski płynące ze współpracy grupowej, konsultacji oraz literatury naukowej. Prace studenta cechuje wysoka dbałość o szczegóły.

10 MACIERZ REALIZACJI PRZEDMIOTU

EFEKT KSZTAŁCENIA	ODNIESIENIE DANEGO EFEKTU DO SZCZEGÓŁOWYCH EFEKTÓW ZDEFINIOWANYCH DLA PROGRAMU	CELE PRZEDMIOTU	TREŚCI PROGRAMOWE	NARZĘDZIA DYDAKTYCZNE	SPOSOBY OCENY
EK1	I2_W01 I2_W02 I2_W06	Cel 1	W1 W2 W3 W4 W5 W6 W7 W8 W9 W10 W11 W12	N1 N2 N5 N6 N7	F1 F2 F3 P1
EK2	I2_U01b I2_U02b I2_U04b I2_U07 I2_U12	Cel 2	L1 L2 L3 L4 L5 L6 L7 L8 L9 L10 L11 L12 L13 W1 W2 W3 W4 W5 W6 W7 W8 W9 W10 W11 W12	N1 N2 N3 N4 N5 N6 N7	F1 F2 F3 P1
EK3	I2_U01b I2_U02b I2_U04b I2_U07 I2_U12	Cel 2	L1 L2 L3 L4 L5 L6 L7 L8 L9 L10 L11 L12 L13 W1 W2 W3 W4 W5 W6 W7 W8 W9 W10 W11 W12	N1 N2 N3 N4 N5 N6 N7	F1 F2 F3 P1

EFEKT KSZTAŁCENIA	ODNIESIENIE DANEGO EFEKTU DO SZCZEGÓŁOWYCH EFEKTÓW ZDEFINIOWANYCH DLA PROGRAMU	CELE PRZEDMIOTU	TREŚCI PROGRAMOWE	NARZĘDZIA DYDAKTYCZNE	SPOSOBY OCENY
EK4	I2_K01 I2_K04	Cel 3	L1 L2 L3 L4 L5 L6 L7 L8 L9 L10 L11 L12 L13	N2 N3 N4 N5 N6 N7	F1 F2 F3 P1

11 WYKAZ LITERATURY

LITERATURA PODSTAWOWA

- [1] | Garrett Golemund, Hadley Wickham, R for Data Science, 2016, O'Reilly, książka dostępna w Internecie na str.: <http://r4ds.had.co.nz/> ; dodatkowe tutoriały i kody są dostępne na stronach: <https://github.com/tidyverse> <http://tidyverse.org> <https://github.com/hadley/r4ds> <https://github.com/hadley/dplyr> <https://github.com/tidyverse/ggplot2>
- [2] | Y. Zhao, "R and Data Mining: Examples and Case Studies", 2014, książka dostępna w WWW z licznymi materiałami, np. książką: <http://www.rdatamining.com/docs/introduction-to-data-mining-with-r> oraz przykłady w R: <http://www.rdatamining.com/examples> oraz wiele inn. wartościowych plików: <http://www.rdatamining.com>)
- [3] | Liczne przykłady gotowych internetowych Notebooków wraz z tutorialami i możliwość tworzenia nowych (swoich) online w środowisku programistycznym RStudio: <https://rnotebook.io/>, <https://rstudio.cloud/> , <https://spark.rstudio.com/> oraz pod adresem: <https://rpubs.com> (lub w ekosystemie SPARK w Chmurze z wykorzystaniem języka R: <https://docs.databricks.com/spark/latest/training/index.html>)
- [4] | Online dokumentacja ekosystemu Apache SPARK w Chmurze (stale przez autorów uaktualniana): <http://spark.apache.org/docs/latest/> (oraz online tutorial pakietu R'a SparkR, pozwalający na bezpośrednio z R'a łączenie się z Apache SPARK (bez udziału RStudio): <http://spark.apache.org/docs/latest/sparkr.html>)
- [4] | R.D. Peng, "R programming for data science", Publ. by Leanpub, 2015, książka dostępna w WWW pod adr.: <http://www.cs.upc.edu/robert/teaching/estadistica/rprogramming.pdf>
- [5] | Online tutoriały tworzenia Notebooków w ekosystemie Apache SPARK w Chmurze (stale przez autorów uaktualniane): <https://docs.databricks.com/spark/latest/training/index.html>

LITERATURA UZUPEŁNIAJĄCA

- [1] | G. James, D. Witten, Trevor Hastie, Robert Tibshirani, An Introduction to Statistical Learning with Applications in R, Springer Series in Statistics, 2016, Stanford, CA, książka dostępna w Internecie na stronie: <http://www-bcf.usc.edu/gareth/ISL> (a wszystkie kody źródłowe z książki są dostępne na stronie: <http://www-bcf.usc.edu/gareth/ISL/code.html>); dużo przykładów autora książki jest na jego stronie: <http://web.stanford.edu/hastie/StatLearning/>
- [2] | Materiały w wersji elektronicznej dostarczone studentom na pierwszych laboratoriach.

12 INFORMACJE O NAUCZYCIELACH AKADEMICKICH

OSOBA ODPOWIEDZIALNA ZA KARTĘ

dr Barbara Borowik (kontakt: bborowik@pk.edu.pl)

OSOBY PROWADZĄCE PRZEDMIOT

1 dr Barbara Borowik (kontakt: bborowik@pk.edu.pl)

13 ZATWIERDZENIE KARTY PRZEDMIOTU DO REALIZACJI

(miejscowość, data)

(odpowiedzialny za przedmiot)

(dziekan)

PRZYJMUJĘ DO REALIZACJI (data i podpisy osób prowadzących przedmiot)

.....