

POLITECHNIKA KRAKOWSKA IM. TADEUSZA KOŚCIUSZKI

KARTA PRZEDMIOTU

obowiązuje studentów rozpoczynających studia w roku akademickim 2020/2021

Wydział Inżynierii i Technologii Chemicznej

Kierunek studiów: Biotechnologia

Profil: Ogólnoakademicki

Forma studiów: stacjonarne

Kod kierunku: B

Stopień studiów: I

Specjalności: Biotechnologia Przemysłowa i w Ochronie Środowiska

1 INFORMACJE O PRZEDMIOCIE

NAZWA PRZEDMIOTU	Podstawy bioinformatyki
NAZWA PRZEDMIOTU W JĘZYKU ANGIELSKIM	Principles of bioinformatics
KOD PRZEDMIOTU	WITCh B oIS A5 20/21
KATEGORIA PRZEDMIOTU	Przedmioty ogólne
LICZBA PUNKTÓW ECTS	4.00
SEMESTRY	2

2 RODZAJ ZAJĘĆ, LICZBA GODZIN W PLANIE STUDIÓW

SEMESTR	WYKŁADY	ĆWICZENIA	LABORATORIUM	LABORATORIUM KOMPUTERO- WE	PROJEKT	SEMINARIUM
2	0	0	0	30	0	0

3 CELE PRZEDMIOTU

Cel 1 Zapoznanie studentów z ważniejszymi algorytmami i metodami stosowanymi obecnie do przechowywania, przetwarzania, analizy, modelowania i wizualizacji olbrzymich ilości danych, także danych chemicznych.

Cel 2 Zapoznanie studentów z ważniejszym oprogramowaniem stosowanym do przechowywania, przetwarzania, analizy, modelowania i wizualizacji olbrzymich ilości danych, także danych chemicznych. Studenci zdobędą

praktyczną wiedzę, jak zainstalować RStudio oraz darmowe, gotowe i nowoczesne (2018 i 2019) biblioteki (tzw. pakiety) w języku R, także z dziedziny chemometrycznej oraz zdobędą praktyczne umiejętności, jak wykorzystywać wspomniane biblioteki, zwłaszcza do przetwarzania, analizy, tworzenia modeli i wizualizacji danych (także chemicznych) zarówno na komputerze, jak i online w Chmurze; jak poprawnie interpretować uzyskane z algorytmów wyniki oraz zdobędą praktyczne umiejętności, jak tworzyć wizualizację danych, zwłaszcza za pomocą internetowych notebooków i dashboard'ów udostępnionych przez RStudio dzięki bibliotekom rmarkdown i shiny.

Cel 3 CEL społeczny: Celem będzie nabycie umiejętności samodzielnego poszerzania swojej wiedzy i doskonalenia umiejętności w dziedzinie bioinformatyki, 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 Wymaganie 1. Podstawy obsługi komputera.
- 2 Wymaganie 2. Podstawowa znajomość języka angielskiego.

5 EFEKTY KSZTAŁCENIA

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

EK2 Umiejętności Umiejętności: Student będzie posiadał umiejętność zastosowania do przechowywania, przetwarzania, analizy i wizualizacji olbrzymich ilości danych ważniejszych algorytmów (tzw. funkcji) zawartych w wybranych bibliotekach (pakietach) języka R i środowiska zintegrowanego RStudio.

EK3 Umiejętności Umiejętności: Student będzie potrafił poprzez RStudio łączyć się internetowo zarówno ze SPARK'iem, tj. z uniwersalnym silnikiem dla Big data, jaki i 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 Kompetencje społeczne: Student nabędzie umiejętność pracy w grupie, pracy indywidualnej, samokształcenia, umiejętność komunikacji z nauczycielem i środowiskiem pozauczelnianym w celu popularyzacji i przedstawiania uzyskanych rezultatów w zrozumiały sposób, samodzielnego poszerzania swojej wiedzy i doskonalenia umiejętności w dziedzinie bioinformatyki, także docenienia wiedzy z tego zakresu w kształtowaniu współczesnej informatyki, nabędzie 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 nabędzie 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 KOMPUTEROWE		
LP	TEMATYKA ZAJĘĆ OPIS SZCZEGÓŁOWY BLOKÓW TEMATYCZNYCH	LICZBA GODZIN

LABORATORIUM KOMPUTEROWE		
LP	TEMATYKA ZAJĘĆ OPIS SZCZEGÓŁOWY BLOKÓW TEMATYCZNYCH	LICZBA GODZIN
K1	Treści programowe 1. Instalacja RStudio, R, ekosystemu Apache SPARK, podstawowych pakietów oferowanych przez RStudio oraz przez R. Proste eksperymenty w środowisku RStudio z wykorzystaniem języka R.	2
K2	Treści programowe 2. Instalacja wybranych pakietów spośród zbioru ponad 1741 bibliotek ze zbioru BioconductorR, omówienie ich oraz wywoływanie w RStudio gotowych kodów źródłowych dostępnych online na podanej studentom stronie oraz http://bioconductor.org/ ; a także proste eksperymenty w środowisku RStudio z wykorzystaniem zainstalowanych pakietów.	2
K3	Treści programowe 3. Instalacja bogatej rodziny "tidyverse" złożonej z kilkunastu nowoczesnych pakietów RStudio, omówienie ich 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	4
K4	Treści programowe 4. Instalacja pakietu sparklyr i dplyr oraz eksperymentowanie z funkcjami tych pakietów: <code>select()</code> , <code>filter()</code> , <code>arrange()</code> , <code>rename()</code> , <code>mutate()</code> , <code>group_by()</code> oraz z operatorem pipeline, tj. <code>%>%</code> .	2
K5	Treści programowe 5. Instalacja i eksperymentowanie z pakietami RStudio i języka R - pozwalającymi na tworzenie ważniejszych modeli dla dużych zbiorów danych (takich jak: modele liniowe, modele drzew regresyjnych i klasyfikacyjnych, modele analizy skupień, i innych modeli). Instalowane pakiety - to: <code>rpart</code> , <code>maptree</code> , <code>rattle</code> , <code>party</code> , <code>randomForest</code> , <code>SVM</code> , <code>naiveBayes</code> , <code>boosting</code> , <code>cluster</code> i kilka innych dodatkowych.	6
K6	Treści programowe 6. Instalacja i eksperymentowanie z pakietem ISLR. Eksperymenty będą bazowały na licznych przykładach zawartych w licznych online laboratoryjnych 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/ (Jest to światowej sławy autor współpracujący przy tworzeniu oprogramowania dla ekosystemu Apache SPARK (i platformy H20), uniwersalnego silnika dla Big Data).	4
K7	Treści programowe 7. Instalacja i eksperymentowanie z pakietami języka R - pozwalającymi na tworzenie internetowych notebooków. Instalowane pakiety - to: <code>magrittr</code> , <code>Rmarkdown</code> i kilka innych mniejszych.	2
K8	Treści programowe 8. Instalacja i eksperymentowanie z pakietami języka R - pozwalającymi na tworzenie dashboardów. Instalowany pakiet, to bogaty w swych funkcjach 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/	4

LABORATORIUM KOMPUTEROWE		
LP	TEMATYKA ZAJĘĆ OPIS SZCZEGÓŁOWY BLOKÓW TEMATYCZNYCH	LICZBA GODZIN
K9	Treści programowe 9. Instalacja i eksperymentowanie z pakietami języka R - pozwalającymi na tworzenie grafów zależności istniejących w dużych zbiorach danych. Instalowane pakiety - to: graphframe, igraph, rgl, snowfall, network, tmap i kilka innych dodatkowych.	2
K10	Treści programowe 10. 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, i kilka innych dodatkowych.	2

7 NARZĘDZIA DYDAKTYCZNE

N1 Narzędzie 1. Ćwiczenia laboratoryjne

N2 Narzędzie 2. Prezentacje multimedialne

N3 Narzędzie 3. Konsultacje

N4 Narzędzie 4. Dyskusja

N5 Narzędzie 5. Praca w 2-3 osobowych grupkach

N6 Narzędzie 6. 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	30
Konsultacje przedmiotowe	20
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	20
Opracowanie wyników	10
Przygotowanie raportu, projektu, prezentacji, dyskusji	40
SUMARYCZNA LICZBA GODZIN DLA PRZEDMIOTU WYNIKAJĄCA Z CAŁEGO NAKŁADU PRACY STUDENTA	120
SUMARYCZNA LICZBA PUNKTÓW ECTS DLA PRZEDMIOTU	4.00

9 SPOSOBY OCENY

OCENA FORMUJĄCA

F1 Cwiczenia praktyczne podczas zajęć laboratoryjnych

F2 Odpowiedzi ustne podczas zajęć laboratoryjnych

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

OCENA PODSUMOWUJĄCA

P1 Średnia ważona ocen formujących

WARUNKI ZALICZENIA PRZEDMIOTU

W1 Warunkiem otrzymania zaliczenia z przedmiotu "Podstawy Bioinformatyki" 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.

OCENA AKTYWNOŚCI BEZ UDZIAŁU NAUCZYCIELA

B1 Osoby mające mało punktów będą mogły podwyższyć sobie punktację wykonując ekstra zlecone sprawozdania.

KRYTERIA OCENY

EFEKT KSZTAŁCENIA 1	
NA OCENĘ 2.0	Brak opanowania zagadnień związanych z Efektem kształcenia 1 i poruszanych na zajęciach w stopniu powyżej 50%.
NA OCENĘ 3.0	Opanowanie zagadnień związanych z Efektem kształcenia 1 i poruszanych na zajęciach w stopniu powyżej 50%, ale poniżej 60%.
NA OCENĘ 3.5	Opanowanie zagadnień związanych z Efektem kształcenia 1 i poruszanych na zajęciach w stopniu powyżej 60%, ale poniżej 70%.
NA OCENĘ 4.0	Opanowanie zagadnień związanych z Efektem kształcenia 1 i poruszanych na zajęciach w stopniu powyżej 70%, ale poniżej 80%.
NA OCENĘ 4.5	Opanowanie zagadnień związanych z Efektem kształcenia 1 i poruszanych na zajęciach w stopniu powyżej 80%, ale poniżej 90%.
NA OCENĘ 5.0	Opanowanie zagadnień związanych z Efektem kształcenia 1 w stopniu powyżej 90%.
EFEKT KSZTAŁCENIA 2	
NA OCENĘ 2.0	Brak opanowania zagadnień związanych z Efektem kształcenia 2 i poruszanych na zajęciach w stopniu powyżej 50%.
NA OCENĘ 3.0	Opanowanie zagadnień związanych z Efektem kształcenia 2 i poruszanych na zajęciach w stopniu powyżej 50%, ale poniżej 60%.
NA OCENĘ 3.5	Opanowanie zagadnień związanych z Efektem kształcenia 2 i poruszanych na zajęciach w stopniu powyżej 60%, ale poniżej 70%.
NA OCENĘ 4.0	Opanowanie zagadnień związanych z Efektem kształcenia 2 i poruszanych na zajęciach w stopniu powyżej 70%, ale poniżej 80%.
NA OCENĘ 4.5	Opanowanie zagadnień związanych z Efektem kształcenia 2 i poruszanych na zajęciach w stopniu powyżej 80%, ale poniżej 90%.
NA OCENĘ 5.0	Opanowanie zagadnień związanych z Efektem kształcenia 2 i poruszanych na zajęciach w stopniu powyżej 90%.
EFEKT KSZTAŁCENIA 3	
NA OCENĘ 2.0	Brak opanowania zagadnień związanych z Efektem kształcenia 3 i poruszanych na zajęciach w stopniu powyżej 50%.
NA OCENĘ 3.0	Opanowanie zagadnień związanych z Efektem kształcenia 3 i poruszanych na zajęciach w stopniu powyżej 50%, ale poniżej 60%.
NA OCENĘ 3.5	Opanowanie zagadnień związanych z Efektem kształcenia 3 i poruszanych na zajęciach w stopniu powyżej 60%, ale poniżej 70%.
NA OCENĘ 4.0	Opanowanie zagadnień związanych z Efektem kształcenia 3 i poruszanych na zajęciach w stopniu powyżej 70%, ale poniżej 80%.
NA OCENĘ 4.5	Opanowanie zagadnień związanych z Efektem kształcenia 3 i poruszanych na zajęciach w stopniu powyżej 80%, ale poniżej 90%.

NA OCENĘ 5.0	Opanowanie zagadnień związanych z Efektem kształcenia 3 i poruszanych na zajęciach 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 swej pracy.
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	K1_W01 K1_W15 b	Cel 1 Cel 2 Cel 3	K1 K2 K3 K4 K5 K6 K7 K8 K9 K10	N1 N2 N3 N4 N5 N6	F1 F2 F3 P1

EFEKT KSZTAŁCENIA	ODNIESIENIE DANEGO EFEKTU DO SZCZEGÓŁOWYCH EFEKTÓW ZDEFINIOWANYCH DLA PROGRAMU	CELE PRZEDMIOTU	TREŚCI PROGRAMOWE	NARZĘDZIA DYDAKTYCZNE	SPOSOBY OCENY
EK2	K1_U01 b K1_U02 K1_U05 K1_U07 K1_U09 b K1_U10 b K1_U11 K1_U15 b	Cel 1 Cel 2 Cel 3	K1 K2 K3 K4 K5 K6 K7 K8 K9 K10	N1 N5 N6	F1 F2 F3 P1
EK3	K1_U01 b K1_U02 K1_U05 K1_U07 K1_U09 b K1_U10 b K1_U11 K1_U15 b	Cel 1 Cel 2 Cel 3	K1 K2 K3 K4 K5 K6 K7 K8 K9 K10	N1 N5 N6	F1 F2 F3 P1
EK4	K1_K01 K1_K02 K1_K03 K1_K04 K1_K05	Cel 1 Cel 2 Cel 3	K1 K2 K3 K4 K5 K6 K7 K8 K9 K10	N1 N5 N6	F1 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] | 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/StatLearn/Books/ISL/>
- [3] | Roger D. Peng, R programming for data science, 2015, Wyd. Leanpub; książka dostępna w Internecie na str.: <http://www.cs.upc.edu/robert/teaching/estadistica/rprogramming.pdf>
- [4] | Liczne przykłady gotowych Notebooków wraz z tutorialami i możliwość tworzenia nowych (swoich) online w ekosystemie SPARK w Chmurze z wykorzystaniem języka R i środowiska RStudio: <https://rnotebook.io/>; <https://rstudio.cloud/> oraz pod adresem: <https://rpubs.com>
- [5] | Liczne przykłady wprowadzające do języka R i ukazujące przykłady eksperymentów (wraz z dokładnymi ich omówieniem) wykonanych w środowisku RStudio i R z wykorzystaniem zbioru 1741 pakietów Bioconductor

- dostępne są na stronie: <http://bcb.dfci.harvard.edu/aedin/courses/BiocDec2011/> (Introduction to R and Bioconductor).

[6] Własne materiały - które zostaną dostarczone studentom na pierwszych zajęciach w wersji elektronicznej.

LITERATURA UZUPEŁNIAJĄCA

[1] Per Bruun Brockhoff, R in 27411: Biological Data Analysis and Chemometrics, 2014, Tutorial dostępny w Internecie na stronie: <http://www2.imm.dtu.dk/courses/27411/Rin27411.pdf>

[2] Heide Garcia and Peter Filzmoser, Multivariate Statistical Analysis using the R package chemometrics, 2017, Vienna University of Technology, Austria, (71 str.), Tutorial dostępny w Internecie na stronie: <https://cran.r-project.org/web/packages/chemometrics/vignettes/chemometrics-vignette.pdf>

[3] CHEMOMETRICS, 10 tutorialów dostępnych w Internecie z mnóstwem przykładów i eksperymentów w języku R w środowisku RStudio. Tych 10 adresów internetowych zostanie podanych na pierwszych zajęciach.

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)

.....