



KARTA OPISU PRZEDMIOTU - SYLABUS

Nazwa przedmiotu

Analiza danych genomicznych w środowisku Bioconductor [S2Bioinf2>BIOCON]

Przedmiot

Kierunek studiów
Bioinformatyka

Rok/Semestr
2/3

Studia w zakresie (specjalność)
–

Profil studiów
ogólnoakademicki

Poziom studiów
drugiego stopnia

Język oferowanego przedmiotu
polski

Forma studiów
stacjonarne

Wymagalność
obieralny

Liczba godzin

Wykład
30

Laboratorium
30

Inne (np. online)
0

Ćwiczenia
0

Projekty/seminaria
0

Liczba punktów ECTS

4,00

Koordynatorzy

dr hab. inż. Aleksandra Świercz
aleksandra.swiercz@put.poznan.pl

Wykładowcy

Wymagania wstępne

Osoba podejmująca studia na II stopniu Bioinformatyki powinna mieć osiągnięte efekty kształcenia z I stopnia tego kierunku studiów, zdefiniowane w Uchwale Senatu PP - efekty te prezentowane są w serwisie internetowym wydziału <http://www.cat.put.poznan.pl/> Student rozpoczynający ten moduł powinien posiadać podstawową wiedzę z zakresu biologii molekularnej, genomiki i transkryptomiki, a także programowania, statystyki i analizy bioinformatycznej sekwencji biologicznych. Powinien posiadać umiejętność rozwiązywania podstawowych problemów biologicznych i bioinformatycznych, testowania i poprawiania błędów w zaimplementowanych przez siebie programach oraz pozyskiwania informacji ze wskazanych źródeł i korzystania z baz danych. Ponadto w zakresie kompetencji społecznych student musi prezentować takie postawy, jak uczciwość, odpowiedzialność, wytrwałość, ciekawość poznawcza, kreatywność, kultura osobista, szacunek dla innych ludzi.

Cel przedmiotu

1. Przekazanie studentom podstawowej wiedzy na temat środowiska Bioconductor, a w szczególności: struktury i typów danych, dostępnych pakietów, możliwości analizy i graficznej wizualizacji danych 2. Zapoznanie studentów z aktualnym zestawem pakietów Bioconductor, ich przeznaczeniem oraz przykładami wykorzystania, z ukierunkowaniem na omówienie podejść analitycznych w genomice 3. Rozwijanie u studentów umiejętności rozwiązywania problemów z zakresu analizy danych genomicznych w środowisku Bioconductor, wyboru i zastosowania narzędzi oraz pakietów Bioconductor do analizy wybranych typów danych biologicznych, a także zastosowania graficznej wizualizacji danych do wspomagania analizy i interpretacji wyników.

Przedmiotowe efekty uczenia się

Wiedza:

1. Ma uporządkowaną, podbudowaną teoretycznie wiedzę w zakresie środowiska Bioconductor oraz zastosowania wybranych pakietów do analiz w skali genomowej i na poziomie populacji
2. Ma wiedzę o zastosowaniach pakietów Bioconductor do analizy bioinformatycznej i integracji różnych typów danych genomicznych oraz wizualizacji danych biologicznych
3. Zna zastosowania pakietów Bioconductor do analizy danych genetycznych w medycynie
4. Zna metody, techniki i narzędzia wykorzystywane w procesie rozwiązywania złożonych zadań bioinformatycznych, głównie o charakterze inżynierskim
5. Zna zasady planowania badań z zakresu bioinformatyki

Umiejętności:

1. Potrafi pozyskać dane źródłowe opisane w publikacjach naukowych (np. wyniki analiz genomicznych) oraz w sposób zintegrowany pozyskiwać z różnych źródeł i interpretować informacje na ich temat (np. adnotacje funkcjonalne, ścieżki powiązań)
2. Potrafi wskazać odpowiedni pakiet środowiska Bioconductor do typowych analiz danych genomicznych oraz zastosować je w praktyce
3. Potrafi stosować zaawansowane techniki i narzędzia informatyczne do rozwiązywania problemów biologicznych oraz ocenić ich przydatność
4. Wykorzystuje środowisko Bioconductor do przetworzenia danych genomicznych i statystycznej analizy wyników
5. Przygotowuje w języku polskim i angielskim prezentację wyników prac badawczych, dyskutuje uzyskane wyniki w kontekście istniejącej wiedzy naukowej
6. Formułuje i testuje hipotezy związane z problemami bioinformatycznymi.

Kompetencje społeczne:

1. Rozwija postawę otwartości na współpracę w środowisku akademickim i przemysłowym, działając na rzecz przestrzegania zasad etyki i dobrych praktyk w analizie danych biologicznych i biomedycznych.
2. Jest świadomy znaczenia ochrony danych wrażliwych i przestrzega zasad związanych z ich przetwarzaniem, w szczególności w kontekście danych genomicznych, które mogą mieć bezpośredni wpływ na zdrowie i prywatność pacjentów.

Metody weryfikacji efektów uczenia się i kryteria oceny

Efekty uczenia się przedstawione wyżej weryfikowane są w następujący sposób:

Ocena formująca

a) w zakresie wykładów weryfikowanie założonych efektów kształcenia realizowane jest przez:

- na podstawie aktywności w dyskusji na temat omawianego materiału;

b) w zakresie laboratoriów weryfikowanie założonych efektów kształcenia realizowane jest przez:

- na podstawie oceny bieżącego postępu realizacji zadań;

Ocena podsumowująca

a) w zakresie wykładów weryfikowanie założonych efektów kształcenia realizowane jest przez:

- kolokwium pisemne składające się z 5 pytań / zadań problemowych - każde zadanie punktowane 0-4 pkt (zadania mogą składać się z kilku podpunktów - za każdy podpunkt jest wówczas wyznaczona punktacja częściowa). Aby uzyskać zaliczenie należy zdobyć co najmniej 11 punktów.

b) w zakresie laboratoriów / ćwiczeń weryfikowanie założonych efektów kształcenia realizowane jest przez:

- ocena końcowa stanowi średnią oceny z opracowania i przedstawienia wybranego pakietu / grupy pakietów oraz ocen za wykonanie poszczególnych ćwiczeń praktycznych z zakresu analizy danych w

Bioconductor w trakcie semestru. Za przedstawienie wybranego pakietu można otrzymać maksymalnie 10 punktów. Za każde ćwiczenie / raport (8-10 ćwiczeń) można otrzymać maksymalnie 5 punktów. Zaliczenie można otrzymać po uzyskaniu więcej niż 50% sumarycznej liczby punktów pod warunkiem złożenia wszystkich wymaganych raportów.

Treści programowe

Przedmiot poświęcony jest teoretycznemu i praktycznemu przedstawieniu pakietu Bioconductor, dedykowanego analizie oraz wizualizacji danych biologicznych, dostępnego bezpłatnie na otwartej licencji GPL. W ramach wykładów omówione zostaną następujące zagadnienia: wprowadzenie do Bioconductor / utrwalenie wiedzy nt. typów zmiennych i operacjach na nich; wczytywania i zapisywania danych z/ do różnych formatów i programów zewnętrznych; przetwarzania danych i tworzenia raportów; integracji z innymi programami, wybranych procedur statystycznych w R; przegląd i instalacja dostępnych edytorów przegląd przydatnych zasobów internetowych; omówienie aktualnego zasobu dostępnych pakietów i ich zastosowań, poznanie i praca z wybranymi bibliotekami umożliwiającymi analizę i wizualizację danych genomicznych. Zagadnienia będą ilustrowane przykładami z literatury naukowej. Laboratoria polegają na analizie prac naukowych oraz na analizie danych biologicznych i biomedycznych, z nastawieniem na poznanie możliwości pakietów omawianych w części wykładowej i ich zastosowań. Zakres omawianych pakietów jest aktualizowany tak, aby uwzględniać najnowsze osiągnięcia i trendy w genomice. Ponadto każdy student wybiera, opracowuje i prezentuje pakiet / grupę pakietów oraz jego możliwości, wg scenariusza uzgodnionego z prowadzącym przedmiot.

Tematyka zajęć

Zapoznanie się ze środowiskiem R oraz z dostępem i różnorodną tematyką pakietów Bioconductor. Przygotowanie wykresów podsumowujących analizy bioinformatyczne, gotowych do publikacji z wykorzystaniem pakietu ggplot2.

Omówienie wybranych pakietów Bioconductor.

Przygotowanie prezentacji wybranych pakietów, według zainteresowań studentów.

Metody dydaktyczne

Wykład ilustrowany prezentacją multimedialną zawierającą omawiane treści programowe, wzbogacone przykładami;

Laboratoria: ćwiczenia praktyczne z zakresu analizy danych w Bioconductor, praca z pakietami R, praca grupowa, dyskusja i analiza problemów

Literatura

Podstawowa:

Biecek P. Przetwarzanie danych w programie R; epub; <http://biecek.pl/R/>

Biecek P. Wizualizacja i modelowanie; epub; <http://biecek.pl/R/>

Uzupełniająca:

1. Biecek P. Przewodnik po pakiecie R. Oficyna Wydawnicza GiS 2017 (wydanie 4), ISBN 978-83-89020-79-6

Górecki T. Podstawy statystyki z przykładami w R. Wydawnictwo BTC 2011, ISBN 978-83-60233-69-6.

Bilans nakładu pracy przeciętnego studenta

	Godzin	ECTS
Łączny nakład pracy	105	4,00
Zajęcia wymagające bezpośredniego kontaktu z nauczycielem	60	2,50
Praca własna studenta (studia literaturowe, przygotowanie do zajęć laboratoryjnych/ćwiczeń, przygotowanie do kolokwium/egzaminu, wykonanie projektu)	45	1,50