



## KARTA OPISU PRZEDMIOTU - SYLABUS

Nazwa przedmiotu

Analiza filogenetyczna [S2Bioinf2>AFIL]

### Przedmiot

Kierunek studiów

Bioinformatyka

Rok/Semestr

1/2

Studia w zakresie (specjalność)

–

Profil studiów

ogólnoakademicki

Poziom studiów

drugiego stopnia

Język oferowanego przedmiotu

polski

Forma studiów

stacjonarne

Wymagalność

obligatoryjny

### Liczba godzin

Wykład

15

Laboratorium

15

Inne

0

Ćwiczenia

0

Projekty/seminaria

0

### Liczba punktów ECTS

2,00

### Koordynatorzy

mgr Mateusz Twardawa

mateusz.twardawa@put.poznan.pl

### Wykładowcy

### Wymagania wstępne

Od studentów oczekuje się znajomości procesów ewolucyjnych i biologii molekularnej, podstawowej znajomości algorytmiki i struktur danych, uczenia maszynowego, metod optymalizacji i statystycznej analizy danych.

### Cel przedmiotu

Celem przedmiotu jest przekazanie studentom wiedzy z zakresu filogenetyki i algorytmów służących do rekonstrukcji drzew filogenetycznych. Studenci w ramach przedmiotu nabędą praktyczne umiejętności posługiwania się metodami i narzędziami pozwalającymi na budowę drzew filogenetycznych i analizę statystyczną uzyskanych wyników.

### Przedmiotowe efekty uczenia się

Wiedza:

1. Student posiada wiedzę na temat filogenetyki i potrafi ją wykorzystywać do przeprowadzenia analizy filogenetycznej opartej o dane biologiczne.
2. Student zna i rozumie algorytmy i metody wykorzystywane w analizie filogenetycznej i posiada wiedzę z zakresu analizy statystycznej drzew filogenetycznych.

3. Student posiada wiedzę o wybranych narzędziach informatycznych i algorytmach do analizy filogenetycznej. Student rozumie techniki filogenetyczne i problemy badawcze w analizie podobieństwa sekwencji biologicznych.
4. Student zna zaawansowane pojęcia z zakresu analizy filogenetycznej i potrafi przeprowadzić analizę filogenetyczną różnych organizmów wykorzystując to tego wiedzę związaną z ewolucją molekularną genomów. Student posiada wiedzę potrzebną do filogenetycznej analizy grup organizmów.
5. Student zna bazy danych, narzędzia i współcześnie stosowane metody w analizie filogenetycznej i rozumie trendy rozwojowe w tej dziedzinie badań.

#### Umiejętności:

1. Student potrafi posługiwać się wybranymi narzędziami informatycznymi do analizy filogenetycznej. Student rozumie techniki filogenetyczne i potrafi je wykorzystać w analizie podobieństwa sekwencji biologicznych. Student potrafi zastosować odpowiednie narzędzia, techniki i metody filogenetyczne do rozwiązywania problemów biologicznych w odniesieniu do filogenezy (w zakresie omawianym na zajęciach).
2. Student potrafi korzystać z metod i narzędzi statystycznych do analizy i oceny drzew filogenetycznych.
3. Student potrafi przeanalizować problem z zakresu filogenezy i dobrać metody do jego rozwiązania.
4. Student potrafi formułować hipotezy filogenetyczne i przeprowadzić odpowiednie analizy aby je przetestować.

#### Kompetencje społeczne:

1. Rozumie potrzebę systematycznego aktualizowania swojej wiedzy z zakresu biologii i informatyki oraz dostrzegania możliwości jej praktycznego zastosowania (w trakcie zajęć będą też omawiane zagadnienia związane wykorzystaniem analizy filogenetycznej do monitorowania ewolucji wirusów podczas epidemii oraz wybrane informacje z najnowszych publikacji związanych z algorytmami do analizy filogenetycznej).

### Metody weryfikacji efektów uczenia się i kryteria oceny

Efekty uczenia się przedstawione wyżej weryfikowane są w następujący sposób:

#### Ocena formująca:

W zakresie wykładów weryfikowanie założonych efektów kształcenia realizowane jest przez:

- odpowiedzi na pytania dotyczące materiału omówionego na poprzednich wykładach,

W zakresie laboratorium weryfikowanie założonych efektów kształcenia realizowane jest przez:

- ocenę przygotowania studenta do poszczególnych laboratoriów oraz ocenę umiejętności związanych z realizacją zaplanowanych zadań,
- ocenianie ciągle, na każdych zajęciach (odpowiedzi ustne) - premiowanie przyrostu umiejętności posługiwania się poznanymi zasadami i metodami,
- dominującym składnikiem oceny formującej stanowią oceny sprawozdań przygotowywanych częściowo w trakcie zajęć, a częściowo po ich zakończeniu;

#### Ocena podsumowująca:

Wiedza zdobyta podczas wykładów zostanie zweryfikowana za pomocą testu.

W zakresie laboratorium weryfikowanie założonych efektów kształcenia realizowane jest przez ocenę sprawozdań realizowanych w ramach poszczególnych zajęć laboratoryjnych. Każdy zbiór zadań jest punktowany. Należy zdobyć ponad 50% punktów przypisanych do laboratorium, aby je zaliczyć. Należy zaliczyć wszystkie sprawozdania, aby otrzymać pozytywną ocenę na koniec zajęć z ćwiczeń. Ocena końcowa z laboratorium jest oparta na łącznej liczbie zdobytych punktów ze wszystkich zajęć.

Na całym przedmiocie obowiązują standardowe progi procentowe: Poniżej 51%. - ocena niedostateczna (2,0); od 51 do 60 %; ocena dostateczna (3,0); od 61 do 70 %. - ocena dostateczna plus (3,5); od 71 do 80 %. - ocena dobra (4,0); od 81 do 90 %. - ocena dobra plus (4,5); od 91 do 100 %. - ocena bardzo dobra (5,0).

### Treści programowe

W ramach pierwszej części wykładów studenci zostaną wprowadzeni do analizy filogenetycznej. Podczas tej części zostaną omówione problemy i potrzeby jakie są adresowane przez analizę filogenetyczną. Zostanie też przedstawiony zarys systematyki prokariotów i eukariotów oraz rolę jaką ma systematyka, taksonomia i analiza filogenetyczna we współczesnych badaniach biologicznych i medycznych.

Druga część wykładu będzie skoncentrowana na molekularnych mechanizmach procesów ewolucyjnej oraz zmienności genomów. Zostaną przedstawione problemy z zakresu analizy filogenetycznej

organizmów wynikające z błędnego wykorzystania danych biologicznych oraz trudności w rekonstrukcji filogenezy w oparciu o poprawne dane. W ramach tej części zostaną też omówione pojęcia takie jak homoplazja i homologia, a także wykorzystanie cech morfologicznych i innych do wsparcia analizy filogenetycznej. Dodatkowo zostaną też zaprezentowane metody i kryteria doboru cech i sekwencji biologicznych do analizy filogenetycznej blisko i daleko spokrewnionych taksonów (markery filogenetyczne). W drugiej części zostanie też omówione pojęcie zegara molekularnego i problemy związane z tym zagadnieniem.

Trzecia część wykładów będzie omawiać różne rodzaje drzew filogenetycznych i grafowych reprezentacji podobieństwa. Zostaną omówione metody transformacji i obróbki danych do analizy filogenetycznej, naturalne i sztuczne taksony, a także ukorzenienie drzewa filogenetycznego.

Czwarta część wykładów będzie poświęcona algorytmom oraz narzędziom do rekonstrukcji drzew filogenetycznych. Studenci zostaną wprowadzeni do problemu grupowania i etykietowania danych (klasteryzacji). Po zarysowaniu ogólnych pojęć, zostaną zaprezentowane w szczególności wiodące metody analizy filogenetycznej tj. metody odległościowe, maksymalnej parsymonii, maksymalnej wiarygodności oraz metody bayesowskie. Na wykładzie zostaną też krótko scharakteryzowane narzędzia, które operują omawianymi algorytmami.

Podczas piątej części wykładów zostaną zaprezentowane metody analizy statystycznej drzew filogenetycznych i wyników analizy filogenetycznej, metody próbkowania drzew i szacowania wewnętrznego wsparcia gałęzi. Zostaną omówione takie zagadnienia jak sygnał filogenetyczny, indeks Bremera, bootstrap, aLRT, czy prawdopodobieństwo a posteriori. Dodatkowo zostaną pokazane metody analizy czasu dywergencji i tworzenia drzew konsensusowych.

Szósta część wykładów będzie skupiona na zastosowaniu heurystyk do obliczeń filogenetycznych oraz na wizualizacji i opracowywaniu wyników analizy filogenetycznej. W ramach tej części zostaną też przedstawione metody oceny i walidacji drzew filogenetycznych.

Siódma część wykładów zostanie oparta o aplikacje analizy filogenetycznej do badań biologicznych. Zostaną przedstawione metody analizy filogenetycznej białek, a także analiza i monitorowanie ewolucji wirusów podczas epidemii. Dodatkowo zostaną przedstawione współczesne trendy oraz potrzeby związane z filogenetyką.

Podczas laboratorium studenci będą rozwiązywać zadania i problemy filogenetyczne, co pozwoli na wykorzystanie w praktyce wiedzy zdobytej na zajęciach. Każde laboratorium będzie polegać na pracy z narzędziami informatycznymi i bazami danych. Na laboratorium studenci będą budować, wizualizować oraz analizować drzewa filogenetyczne za pomocą różnych metod. Jednym z tematów poruszanych na zajęciach laboratoryjnych będzie monitorowanie ewolucji wirusów, co jest praktycznym wykorzystaniem metod filogenetycznych w zapobieganiu i analizie rozwoju epidemii.

## Tematyka zajęć

Wykład:

1. Wprowadzenie do filogenetyki
2. Modele ewolucji molekularnej i obliczanie dystansu dla sekwencji biologicznych
3. Datowanie molekularne i drzewa filogenetyczne
4. Przeszukiwanie drzew i metody oparte na analizie prawdopodobieństwa
5. Wizualizacja, analiza i ocena otrzymanych drzew
6. Zastosowania i współczesne kierunki rozwoju filogenetyki

Laboratoria:

1. Tworzenie drzew w oparciu o dopasowania wielu sekwencji
2. Metody dystansowe - metody addytywne i macierze dystansu
3. Metody dystansowe - drzewa ultrametryczne i UPGMA
4. Metody dystansowe - algorytm Neighbor-Joining
5. Największa parsymonia i największa wiarygodność
6. Metody Bayesowskie w analizie filogenetycznej
7. Wizualizacja, analiza i ocena otrzymanych drzew
8. Monitorowanie ewolucji patogenów

## Metody dydaktyczne

1. Wykład: prezentacja multimedialna, prezentacja ilustrowana przykładami podawanymi na tablicy.
2. Laboratorium:
  - wykonywanie zadań praktycznych (w tym związanych z analizą i obróbką surowych danych, statystyką, wizualizacją i ewaluacją otrzymywanych drzew),

- implementacja i walidacja studenckich wersji wybranych algorytmów filogenetycznych
- wykorzystywanie do pracy popularnego oprogramowania do analizy filogenetycznej z jednoczesnym omawianiem działania wybranych metod analitycznych oraz algorytmów wykorzystywanych do tworzenia, wizualizacji i oceny drzew filogenetycznych
- opracowywanie wyników w formie raportów,
- dyskusja podczas zajęć.

## Literatura

Podstawowa:

Hall B., Łatwe drzewa filogenetyczne, WUW, Warszawa, 2008

Uzupełniająca:

Felsenstein J., Inferring phylogenies, Sinauer Associates, , 2004

Warnow, T., Computational Phylogenetics: An Introduction to Designing Methods for Phylogeny Estimation. Cambridge, Cambridge University Press, 2017

Xia, X. A Mathematical Primer of Molecular Phylogenetics, Apple Academic Press Inc., 2020

## Bilans nakładu pracy przeciętnego studenta

|                                                                                                                                                      | Godzin | ECTS |
|------------------------------------------------------------------------------------------------------------------------------------------------------|--------|------|
| Łączny nakład pracy                                                                                                                                  | 55     | 2,00 |
| Zajęcia wymagające bezpośredniego kontaktu z nauczycielem                                                                                            | 30     | 1,00 |
| Praca własna studenta (studia literaturowe, przygotowanie do zajęć laboratoryjnych/ćwiczeń, przygotowanie do kolokwium/egzaminu, wykonanie projektu) | 25     | 1,00 |