



KARTA OPISU PRZEDMIOTU - SYLABUS

Nazwa przedmiotu

Bioinformatyka strukturalna RNA [S2Bioinf2>RNA]

Przedmiot

Kierunek studiów
Bioinformatyka

Rok/Semestr
1/1

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
30

Inne
0

Ćwiczenia
0

Projekty/seminaria
0

Liczba punktów ECTS

4,00

Koordynatorzy

prof. dr hab. inż. Marta Szachniuk
marta.szachniuk@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 Informatyki i Telekomunikacji. Student rozpoczynający niniejszy moduł powinien posiadać podstawową wiedzę o biologii strukturalnej, teorii algorytmów, programowaniu, bazach danych oraz algorytmach kombinatorycznych. Powinien posiadać umiejętność rozwiązywania podstawowych problemów z zakresu programowania i analizy danych biologicznych. Ponadto w zakresie kompetencji społecznych student winien prezentować takie postawy jak uczciwość, odpowiedzialność, wytrwałość, ciekawość poznawcza, kreatywność, kultura osobista, szacunek dla innych ludzi.

Cel przedmiotu

1. Przekazanie studentom podstawowej wiedzy o algorytmach i metodach optymalizacji stosowanych do rozwiązywania podstawowych problemów współczesnej biologii, biochemii i bioinformatyki strukturalnej RNA oraz o najlepszych narzędziach umożliwiających magazynowanie, przetwarzanie, analizę i pozyskiwanie danych biologicznych o cząsteczkach kwasów nukleinowych. 2. Zapoznanie studentów z architekturą najpopularniejszych narzędzi stosowanych w bioinformatyce strukturalnej RNA, zarówno prekursorskich, jak i tych stosowanych dziś, w tym najnowszych modeli opartych na uczeniu maszynowym. Nacisk położony jest na wady i zalety stosowanych rozwiązań, ich ograniczenia (których zazwyczaj nie są świadomi użytkownicy tych systemów), wykorzystywane technologie, jakość otrzymywanych rozwiązań, ich wiarygodność oraz sposoby weryfikacji wyników. 3. Rozwinięcie u studentów umiejętności matematycznego modelowania oraz rozwiązywania problemów obliczeniowych w biochemii i biologii strukturalnej RNA przy pomocy zarówno prostych, jak i zaawansowanych algorytmów i modeli bioinformatycznych, umiejętności przygotowywania danych dla modeli głębokiego uczenia i trenowania tych modeli oraz testowania nowych metod obliczeniowych.

Przedmiotowe efekty uczenia się

Wiedza:

1. Ma rozszerzoną i pogłębioną wiedzę na temat modeli matematycznych, statystycznych, optymalizacyjnych oraz modeli sztucznej inteligencji stosowanych do reprezentowania struktury cząsteczek kwasów nukleinowych na różnym poziomie szczegółowości, ich modelowania oraz analizy procesów przebiegających z udziałem tych cząsteczek.
2. Zna metody, techniki i narzędzia wykorzystywane w procesie rozwiązywania złożonych zadań z zakresu analizy struktur RNA, głównie o charakterze inżynierskim.
3. Zna architekturę oraz szczegóły funkcjonalności najpopularniejszych strukturalnych repozytoriów danych oraz algorytmów wspomagających obliczenia strukturalne (przewidywanie struktur, rozwiązywanie struktur, ewaluacja i porównywanie struktur, klastrowanie, etc).
4. Zna zagadnienia z zakresu modelowania i analizy systemów oraz struktur biologicznych oparte na solidnych podstawach teoretycznych.
5. Ma szczegółową wiedzę na temat komputerowego modelowania procesów przewidywania, modelowania, porównywania i ewaluacji struktur molekularnych.
6. Ma opartą na solidnych podstawach teoretycznych szczegółową wiedzę na temat planowania badań, optymalizacji i efektywnych algorytmów wykorzystywanych w bioinformatyce strukturalnej RNA. Zna działanie szeregu najważniejszych historycznie i najpopularniejszych algorytmów do przewidywania i obliczania struktur, znajdowania wspólnych podstruktur, dopasowywania struktur.
7. Wie, w jakim kierunku rozwijają się metody obliczeniowe dla problemów istotnych w biologii i biochemii strukturalnej RNA wymagających wspomaganie algorytmicznego oraz jakie metody proponuje się dla nowych problemów w bioinformatyce strukturalnej.

Umiejętności:

1. Wykorzystuje i interpretuje informacje pozyskane z literatury dotyczącej ogólnych problemów bioinformatyki strukturalnej RNA, jak i źródeł specjalistycznych (publikacji naukowych w czasopismach typu Bioinformatics, BMC Bioinformatics, Nucleic Acids Research, Nature Methods, itp.; serwisów i portali internetowych z dziedziny). Dokonuje interpretacji i krytycznej oceny treści.
2. Stosuje zaawansowane techniki i narzędzia informatyczne do rozwiązywania problemów biologicznych i potrafi ocenić ich przydatność.
3. Stosuje wiedzę z zakresu biochemii i nauk pokrewnych do rozwiązywania problemów bioinformatyki strukturalnej cząsteczek RNA (np. modelowania, adnotowania, oceny struktur).
4. Pod kierunkiem opiekuna naukowego planuje i wykonuje zadania badawcze (np. modelowanie struktury z wykorzystaniem narzędzi bioinformatycznych, ocena jakości modeli RNA, klastrowanie struktur na podstawie ich podobieństw i różnic) z wykorzystaniem poznanych metod obliczeniowych.
5. Stosuje metody matematyczne oraz specjalistyczne techniki i narzędzia informatyczne do reprezentacji, przetwarzania oraz analizy danych strukturalnych opisujących architekturę kwasów nukleinowych na różnych poziomach szczegółowości.
6. Potrafi projektować i tworzyć oprogramowanie komputerowe do rozwiązywania wybranych problemów bioinformatyki strukturalnej używając właściwych metod, technik i narzędzi.
7. Przygotowuje w języku polskim i angielskim prezentację wyników swoich prac badawczych (sprawozdania, prezentacje z wykonania projektu) oraz dyskutuje wyniki tych badań w środowisku naukowym.
8. Potrafi ocenić przydatność nowych osiągnięć w zakresie bioinformatyki i biochemii, zwłaszcza w

przetwarzaniu i analizie danych strukturalnych RNA.

Kompetencje społeczne:

1. Rozumie potrzebę uczenia się przez całe życie (na zajęciach przedstawiane są najnowsze osiągnięcia w dziedzinie bioinformatyki strukturalnej; również student ma możliwość prezentacji tematu dotyczącego najnowszych problemów i ich rozwiązań).
2. Potrafi współdziałać i pracować w grupie (okazją do praktykowania pracy grupowej są zajęcia laboratoryjne z przedmiotu).
3. Rozumie potrzebę systematycznego zapoznawania się z czasopismami naukowymi i popularnonaukowymi w celu poszerzania i pogłębiania wiedzy bioinformatycznej.
4. Wykazuje twórczą postawę w życiu zawodowym i społecznym.

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:

- odpowiedzi na pytania dotyczące materiału omówionego na poprzednich wykładach

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

- ocenę umiejętności związanych z realizacją ćwiczeń laboratoryjnych
- ocenę sprawozdań przygotowywanych częściowo w trakcie zajęć, a częściowo po ich zakończeniu
- ocenę i „obronę” zrealizowanych przez studenta ćwiczeń laboratoryjnych.

Ocena podsumowująca

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

- ocenę wiedzy i umiejętności wykazanych na egzaminie pisemnym, składającym się z 5 pytań (po 3 punkty za pytanie). 20% pytań ma charakter testu wielokrotnego wyboru, a pozostała część ma charakter problemowy. Wszystkie pytania egzaminacyjne dotyczą zagadnień omówionych podczas wykładów oraz zajęć laboratoryjnych. Pozytywne zaliczenie egzaminu jest możliwe jeśli student uzyska łącznie co najmniej 8 punktów.

- omówienie wyników egzaminu

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

- ocenę wiedzy i umiejętności związanych z treściami przekazywanymi na ćwiczeniach poprzez kolokwium końcowe
- zestawienie ocen wystawionych w trakcie semestru w postaci średniej ważonej, w skład której wchodzi oceny cząstkowe wystawiane za wykonane projekty, sprawozdania oraz aktywność wykazaną podczas zajęć.

Aktywność podczas zajęć premiowana jest dodatkowymi punktami, w szczególności za:

- omówienie dodatkowych aspektów zagadnienia,
- przygotowanie krótkiej prezentacji dotyczącej problemu z zakresu bioinformatyki strukturalnej nieprezentowanego przez wykładowcę,
- efektywność zastosowania zdobytej wiedzy podczas rozwiązywania zadanego problemu,
- uwagi prowadzące do udoskonalenia materiałów dydaktycznych lub procesu dydaktycznego.

Treści programowe

Program wykładu obejmuje następujące zagadnienia. Pierwszy wykład stanowi ogólne wprowadzenie do bioinformatyki strukturalnej RNA, w którym przedstawione są kluczowe kwestie współczesnej dziedziny wraz z charakterystyką dostępnych rozwiązań pod kątem wymagań, ograniczeń i złożoności obliczeniowej. Drugi wykład poświęcony jest strukturze drugorzędowej RNA - jej klasyfikacji, reprezentacjom oraz metodom modelowania. Studenci poznają zasady parowania zasad, typowe motywy oraz sposoby zapisu i wizualizacji struktur RNA w różnych formatach. Omawiane są także klasyczne i nowoczesne algorytmy przewidywania struktur drugorzędowych, w tym metody oparte na programowaniu dynamicznym, termodynamice oraz głębokich sieciach neuronowych, wraz z ich zaletami i ograniczeniami. Trzeci wykład wyjaśnia metody eksperymentalne biologii strukturalnej, takie jak krystalografia rentgenowska, spektroskopia NMR i mikroskopia krioelektronowa, oraz komputerowe techniki modelowania trójwymiarowej struktury RNA, w tym metody szablonowe i oparte na uczeniu głębokim. Podkreśla znaczenie jakości danych, ograniczeń metod oraz rolę predykcji w badaniach strukturalnych RNA. Czwarty wykład koncentruje się na wizualizacji i adnotacji struktur 3D RNA, prezentując metody wizualizacji molekularnej, różne reprezentacje strukturalne i formaty danych oraz nowoczesne narzędzia do identyfikacji i wizualizacji pseudowęzłów. W kolejnym wykładzie omawiana

jest walidacja i ocena modeli struktur RNA, ze szczególnym uwzględnieniem zagadnień topologicznych, takich jak węzły i splątania w cząsteczkach RNA. Studenci poznają metody oceny jakości modeli, narzędzia walidacyjne oraz biologiczne i obliczeniowe aspekty splątań. W dalszej części omawiane są algorytmy i miary ilościowe do porównywania struktur na różnych poziomach reprezentacji oraz ich praktyczne zastosowania w ocenie predykcji i identyfikacji motywów strukturalnych. Wykład szósty poświęcony jest metodom oceny dokładności predykcji modeli 3D RNA, obejmującym miary geometryczne, algorytmy optymalnego dopasowania, a także zaawansowane metody analizy kontaktów i topologii, takie jak kąty torsyjne i indeksy deformacji. Kolejny wykład wprowadza metody symulacji Monte Carlo i ich zastosowanie w bioinformatyce strukturalnej, omawiając genezę metody, różnice między symulacjami deterministycznymi a stochastycznymi, oraz przykłady zastosowań, w tym szczegółowy opis algorytmu Metropolis i symulacji na przykładzie programu do modelowania układów makromolekularnych. Podczas następnego wykładu omawiane są podstawy mechaniki i dynamiki molekularnej, w szczególności równania ruchu, funkcje potencjału i siły działające na atomy. Przedstawione są różne pola siłowe służące do opisu oddziaływań molekularnych oraz ich zastosowania w modelowaniu struktur biomolekuł. Wykład obejmuje także wprowadzenie do mechaniki statystycznej oraz podstawowe techniki symulacyjne, takie jak minimalizacja energii, równowaga i próbkowanie, ze szczególnym uwzględnieniem istotnych parametrów symulacji oraz warunków brzegowych. Ostatni wykład przewidziany jest na prezentacje studenckie, podczas których ochotnicy przedstawiają wybrane przez siebie najnowsze problemy i wyzwania bioinformatyki strukturalnej RNA. Ćwiczenia laboratoryjne związane są ściśle z tematyką wykładów. Wykład stanowi wprowadzenie teoretyczne. Podczas ćwiczeń laboratoryjnych studenci uczą się praktycznego wykorzystania narzędzie omawianych podczas wykładu oraz implementują własne metody do rozwiązywania prostych problemów bioinformatycznych. Ćwiczenia laboratoryjne odbywają się w formie dwugodzinnych zajęć w laboratorium komputerowym i mają charakter praktyczny, przygotowując studentów do samodzielnego korzystania z narzędzi i bibliotek programistycznych. Program obejmuje szeroki zakres zagadnień, mocno powiązanych z tematyką wykładów, takich jak zgłębianie kanonicznej struktury drugorzędowej RNA, formaty zapisu oraz definicje motywów strukturalnych. Studenci opracowują konwertery formatów z funkcją heurystycznego rozwiązywania problemu przypisania rzędu pseudowęzłów i narzędzia do identyfikacji motywów. Uczą się przewidywania struktur 2D RNA na podstawie sekwencji i alignmentu, porównywania struktur, analizowania par niekanonicznych oraz wykrywania ewolucyjnie zachowanych modułów 3D RNA. Ćwiczenia obejmują również adnotację struktur drugorzędowych z wykorzystaniem współrzędnych atomowych, klasyczne metody porównywania struktur 3D, a także symulacje dynamiki molekularnej, od zasady uruchamiania symulacji, przez analizę wyników, do realizacji własnych projektów. W programie planowane są również zajęcia z modelowania struktur z użyciem najnowszych modeli sztucznej inteligencji, w tym zadania związane z predykcją trójwymiarowych struktur RNA, prognozowaniem niekanonicznych parowań w strukturze 2D oraz ogólną oceną jakości modeli. Studenci uczą się przygotowywać zbiory danych radząc sobie z problemami redundancji danych i data leakage, a także budują pipeline do automatycznej analizy i filtrowania danych, bazując na rozwiązaniach takich jak RNA3DB. Dodatkowo, ćwiczenia obejmują porównanie metod AI (płytkiego i głębokiego uczenia maszynowego) z metodami klasycznymi w celu oceny skuteczności i wpływu różnych zbiorów testowych na wyniki obliczeń. W ramach zajęć przewidziane są także zadania związane z użyciem głębokich sieci neuronowych do wykrywania konserwatywnych wzorców w miejscach wiązań RNA-białko, mające na celu pogłębienie rozumienia funkcji i struktury tych biomolekuł.

Tematyka zajęć

Tematyka zajęć obejmuje zagadnienia związane z pozyskiwaniem danych o strukturach RNA w drodze eksperymentów obliczeniowych oraz laboratoryjnych, przetwarzaniem danych strukturalnych na różnych poziomach szczegółowości, analizą cech strukturalnych, wizualizacją danych, klasyfikacją topologii, adnotowaniem struktur, wyszukiwaniem motywów, referencyjną i bezreferencyjną oceną struktur, wykrywaniem nieprawidłowości w modelach strukturalnych oraz modelowaniem struktur z udziałem lub bez udziału danych eksperymentalnych. Zajęcia obejmują pracę z metodami klasycznymi oraz modelami głębokiego uczenia.

Metody dydaktyczne

1. Wykład: prezentacja multimedialna, prezentacja ilustrowana przykładami podawanymi na tablicy, pokaz multimedialny.
2. Ćwiczenia laboratoryjne: rozwiązywanie zadań, ćwiczenia praktyczne, wykonywanie eksperymentów obliczeniowych, dyskusja, praca w zespole, pokaz multimedialny, warsztaty, gry integracyjne, studium

przypadków.

Literatura

Podstawowa:

1. J. Gu, P.E. Bourne, "Structural Bioinformatics"
2. P. Baldi, S. Brunak, "Bioinformatics: The Machine Learning Approach"
3. T. Schwede, M. Peitsch, "Computational structural biology. Methods and applications."

Uzupełniająca:

Najnowsze publikacje naukowe z dziedziny bioinformatyki strukturalnej.

Bilans nakładu pracy przeciętnego studenta

	Godzin	ECTS
Łączny nakład pracy	100	4,00
Zajęcia wymagające bezpośredniego kontaktu z nauczycielem	45	2,00
Praca własna studenta (studia literaturowe, przygotowanie do zajęć laboratoryjnych/ćwiczeń, przygotowanie do kolokwium/egzaminu, wykonanie projektu)	55	2,00