**Федеральное государственное автономное образовательное учреждение высшего образования
"Национальный исследовательский университет
"Высшая школа экономики"**

Факультет компьютерных наук

**Рабочая программа дисциплины** «Структурная биоинформатика и моделирование лекарств»

(2 курс)

для образовательной программы «Анализ данных в биологии и медицине» направления подготовки 01.04.02 Прикладная математика и информатика

уровень магистр

Разработчик программы

А.В.Головин, профессор кафедры технологий моделирования сложных систем, golovin@belozersky.msu.ru

Одобрена на заседании кафедры технологий моделирования сложных систем

Зав. Кафедрой А.Н.Соболевский

Утверждена Академическим руководителем программы

Академический руководитель образовательной программы

М.С.Гельфанд

Москва, 2017

*Настоящая программа не может быть использована другими подразделениями университета и другими вузами без разрешения подразделения-разработчика программы.*

1. **Область применения и нормативные ссылки**

Настоящая программа учебной дисциплины устанавливает требования к образовательным результатам и результатам обучения студента и определяет содержание и виды учебных занятий и отчетности.

Программа предназначена для преподавателей, ведущих дисциплину «Структурная биоинформатика и моделирование лекарств», учебных ассистентов и студентов направления подготовки 01.04.02 Прикладная математика и информатика, обучающихся по образовательной программе «Анализ данных в биологии и медицине».

Программа учебной дисциплины разработана в соответствии с:

* Образовательным стандартом НИУ ВШЭ 01.04.02 Прикладная математика и информактика;
* Образовательной программой «Анализ данных в биологии и медицине».
* Объединенным учебным планом университета по образовательной программе «Анализ данных в биологии и медицине», утвержденным в 2017 г.
1. **Цели освоения дисциплины**

Курс посвящен введению в структурную биоинформатику, методы моделирования структуры и динамики биополимеров и низкомолекулярных веществ, основы хемоинформатки, основы квантовой химии и вычислительные методы по разработке биологически активных молекул.

Цели курса:

* познакомить студентов с основными понятиями методами анализа в структурной биоинформатики и методами молекулярного моделирования;
* дать общую картину достижений и проблем современной вычислительной структурной биологии.
* Курс позволит составить остальные курсы модуля в единую и стройную систему знаний.

Целями освоения дисциплины «Структурная биоинформатика и моделирование лекарств» являются ознакомление студентов к подходам моделирования и симуляции сложных молекулярных систем биологического и медицинского значения.

1. **Компетенции обучающегося, формируемые в результате освоения дисциплины**

[Компетенции для программы учебной дисциплины берутся из стандартов: ФГОС/ ОС НИУ ВШЭ для соответствующего уровня и направления подготовки и из числа закрепленных за дисциплиной в матрице компетенций образовательной программы]

Уровни формирования компетенций:

**РБ** — ресурсная база, в основном теоретические и предметные основы (знания, умения);
 **СД** – способы деятельности, составляющие практическое ядро данной компетенции;
 **МЦ** – мотивационно-ценностная составляющая, отражает степень осознания ценности компетенции человеком и готовность ее использовать

В результате освоения дисциплины студент осваивает компетенции:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Компетенция | Код по ОС ВШЭ | Уровень формирования компетенции | Дескрипторы – основные признаки освоения (показатели достижения результата) | Формы и методы обучения, способствующие формированию и развитию компетенции | Форма контроля уровня сформированности компетенции |
| Знания в области структурной биоинформатики и молекулярного моделирования |  |  | Способность владеть фундаментальными понятиями из вычислительной структурной биологии; понимать основные приближения используемые в молекулярном моделировании  | Лекции, научно-исследовательские семинары, домашние задания | Самостоятельное выполнение заданий для практических занятий  |
| Способен анализировать тексты с описанием научно-исследовательских работ |  |  | Способность описывать решение научных задач и формулировать результаты | Лекции, научно-исследовательские семинары, домашние задания | Самостоятельное выполнение заданий для практических занятий  |
| Способен владеть терминологией в области структурной биоинформатики |  |  | Знание основных терминов и теорий в области структурной биоинформатики | Лекции, научно-исследовательские семинары, домашние задания | Самостоятельное выполнение заданий для практических занятий  |
| Способен применять в исследовательской и прикладной деятельности полученные знания при анализе литературы. |  |  | Навык применения полученных знаний при анализе литературы | Лекции, научно-исследовательские семинары, домашние задания | Самостоятельное выполнение заданий для практических занятий  |
| Способность использовать свои знания для формулирования и решения фундаментальных и прикладных задач, связанных с анализом данных структурной биологии |  |  | Способность использовать свои знания для решения фундаментальных и прикладных задач, связанным с анализом молекулярно-биологических данных | Лекции, научно-исследовательские семинары, домашние задания | Самостоятельное выполнение заданий для практических занятий , контрольные работы |

1. **Место дисциплины в структуре образовательной программы**

Настоящая дисциплина относится к блоку дисциплин программы.

Является обязательной.

Изучение данной дисциплины базируется на следующих дисциплинах:

* статистическая физика, органическая химия, молекулярная биология, статистика, анализ данных, программирование.

Для освоения учебной дисциплины студенты должны владеть следующими знаниями и компетенциями:

* Введение в молекулярную биологию: знание основных процессов,
* Курс физики
* Школьный курс органической химии
* Статистический анализ
* Линейная алгебра
* Математический анализ
* Информатика

Основные положения дисциплины должны быть использованы в дальнейшем при изучении дисциплин:

* моделирование сложных процессов
1. **Тематический план учебной дисциплины**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| № | Название раздела | Всего часов  | Аудиторные часы | Самостоятельная работа |
| Лекции | Семинары | Практические занятия | Другие виды работы[[1]](#footnote-0) |
| 1 | Молекулярная механика. Молекулярная динамика. | 70 | 8 | 12 | 50 |  |  |
| 2 | Моделирование структуры белков.Методы предсказания структуры из первых принципов.  | 70 | 8 | 12 | 50 |  |  |
| 3 | Поиск новых био-активных молекул и химоинформатикаМолекулярное моделирование для поиска лекарств. | 44 | 5 | 7 | 32 |  |  |
|  | Молекулярные дескрипторы. Количественное описание структура-активность. Определение и использование 3D фармакофора. | 44 | 5 | 7 | 32 |  |  |
|  | Итого | 228 | 26 | 38 | 164 |  |  |

1. **Формы контроля знаний студентов**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Тип контроля | Форма контроля | 1 год | Кафедра/подразделение | Параметры \*\* |
| 1 | 2 | 3 | 4 |
| Текущий | Контрольная работа | \* | \* |  |  |  | Выполнение контрольного задания , 90 минут |
| Домашнее задание | \* | \* |  |  |  | Обсуждение ключевых компонентов домашней работы |
| Итоговый | Экзамен |  | \* |  |  |  | Устный экзамен |

1. **Критерии оценки знаний, навыков**

В рамках курса студенты должны выполнить два контрольных задания и сдать экзамен, а также выполнять домашние задания. Каждая контрольная работа и экзамен оцениваются по 10-балльной шкале. Домашнее задание состоит в выполнении задания в виде работающего скрипта и оформленных результатах его работы. Экзамен содержит набор теоретических вопросов по различным темам курса.

**Порядок формирования оценки по дисциплине**

***Накопленная оценка*** за текущий контроль формируется из оценок за контрольную работу и домашнее задание следующим образом:

,

где — оценка за контрольные работы.

Каждый студент, сделавший домашнее задание получает оценку по 10-бальной шкале — оценка за домашнее задание, и округляется до целого числа арифметическим способом. При оценке в первом квартиле ряд работ модуля может быть засчитан вместо оценки за одну из контрольных работ.

***Итоговая оценка по дисциплине***  выставляется по десятибалльной шкале, согласно формуле:

,

где — накопленная оценка, — оценка за экзамен, и округляется до целого числа арифметическим способом.

### Таблица соответствия оценок по десятибалльной и пятибалльной системе

|  |  |
| --- | --- |
| По десятибалльной шкале | По пятибалльной системе |
| 1 – неудовлетворительно2 – очень плохо3 – плохо | неудовлетворительно – 2 |
| 4 – удовлетворительно5 – весьма удовлетворительно | удовлетворительно – 3 |
| 6 – хорошо7 – очень хорошо | хорошо – 4 |
| 8 – почти отлично9 – отлично10 – блестяще | отлично – 5 |

1. **Содержание дисциплины**

Тема 1. Поиск новых био-активных молекул и химоинформатика

Молекулярное моделирование для поиска лекарств. Компьютерное представление молекул, SMILES, SMARTS. Источники данных для баз данных структур. Молекулярный докинг. Применение поиска по базам данных структур и докинга. Построение лиганда de novo на основе структуры белка.

Задание для самостоятельной работы: Создать комбинаторную библиотеку и выбрать лучший лиганд.

Тема 2. Квантовая химия

Волновая функция. Уравнение Шредингера. Одно-электронный атом. Базисные наборы.

Задание для самостоятельной работы: Рассчитать первые 5 электронных орбиталей и визуализировать их в PyMol.

Тема 3. Молекулярная механика

Главные особенности молекулярно механических силовых полей. Потенциалы взаимодействий через ковалентные связи. Типы силовых полей. Электростатические взаимодействия. Ван дер Ваальсовы взаимодействия.

Параметры для описания воды. Создание силового поля

Задание для самостоятельной работы: разбор построения поля GAFF

Тема 4. Молекулярная динамика

Простые модели. Создание системы и запуск молекулярной динамики. Динамика с ограничениями. Молекулярная динамика при постоянном давлении и температуре. Введение эффектов растворителя. Конформационные изменения. Гибридый метод QM/MМ. Сканирование фазового пространства, Plumed.

Задание для самостоятельной работы: Сравнение алгоритмов для поддержания постоянной температуры. Моделирование: самосборки липидного бислоя, термической денатурации ДНК, химической денатурации пептида.

Тема 5. Методы Монте-Карло

Расчёт свойств интегрированием. Применение метода Метрополиса Монте-Карло.

Монте-Карло моделирование молекул. «склонные» методы Монте-Карло. Монте-Карло выборка из из различных ансамблей. Что использовать: Монте-Карло или молекулярную динамику?

Задание для самостоятельной работы: Сравнение результатов Монте-Карло и молекулярную динамики.

Тема 6. Моделирование структуры белков.

Методы предсказания структуры из первых принципов. Сравнительное моделирование. Построение и развитие модели. Самосборка и денатурация структуры белка.

Задание для самостоятельной работы: Построить модель для предложенного белка.

Тема 7.

Молекулярный докинг. Молекулярные дескрипторы. Количественное описание структура-активность. Определение и использование 3D фармакофора.

Задание для самостоятельной работы: Предсказание активности низкомолекулярных соединений в отношении данного белка.

1. **Образовательные технологии**

Каждое задание содержит в себе упражнения по подготовке данных для компьютерных симуляций, запуск симуляций и их критический анализ. Работа выполняется в интерактивной среде Jupyter (python)

1. **Оценочные средства для текущего контроля и аттестации студента**
* Вопросы для оценки качества освоения дисциплины
* Волновая функция
* Точное решение уравнения Шредингера Одно-электронный атом
* Метод самосогласованного поля, SCF
* Метод: Хартри-Фока, Подход Рутхана-Хола
* Базисные наборы
* Семи-эмпиричиские методы
* Описания базисных наборов для программы GAUSSIAN
* Теория функционала плотности, DFT
* Молекулярная механика Простое уравнение силового поля
* Электростатические взаимодействия. Двойное обрезание Суммирование Эвальда
* Ван-дер-Ваальсовы взаимодействия Взаимодействия между разными типами атомов
* Минимизация энергии Минимумы, максимумы и стационарные точки Переходные состояния
* Молекулярная динамика Алгоритмы интегратора
* Периодические граничные условия
* Молекулярная динамика Список соседей
* Ограничения быстрых колебаний
* Температура Контроль давления в системе
* Гибридное QM/ММ моделирование
* Метод Монте-Карло Расчёт термодинамических свойств
* Методы выборочного поиска (Biased Monte-Carlo)
* J-walking
* Дистанционная геометрия и ЯМР
* Моделирование отжига
* Использование МД при оптимизации данных ЯМР
* Сравнительное моделирование
* Степень идентичности и сравнительное моделирование
* Предсказание структуры белка Ab initio
* Термодинамическая пертурбация
* Термодинамические циклы
* Потенциал средней силы
* Линейное представление молекул, SMILES
* SMARTS: паттерны для SMILES
* Фрагментарное построение лиганда
* QSAR, количественные соотношения структура/ активность
* Докинг белок-лиганд
* Белковый докинг
	1. **Примеры заданий промежуточной аттестации**

**Билет N1.**

1. Предсказание структуры белка Ab initio
2. Периодические граничные условия

**Билет N2.**

1. QSAR, количественные соотношения структура/ активность
2. Электростатические взаимодействия. Двойное обрезание Суммирование Эвальда
3. **Порядок формирования оценок по дисциплине**

***Накопленная оценка*** за текущий контроль формируется из оценок за контрольную работу и домашнее задание следующим образом:

где — оценка за контрольные работы.

Каждый студент, сделавший домашнее задание получает оценку по 10-бальной шкале — оценка за домашнее задание, и округляется до целого числа арифметическим способом. При оценке в первом квартиле ряд работ модуля может быть засчитан вместо оценки за одну из контрольных работ.

***Итоговая оценка по дисциплине***  выставляется по десятибалльной шкале, согласно формуле:

,

где — накопленная оценка, — оценка за экзамен, и округляется до целого числа арифметическим способом.

### Таблица соответствия оценок по десятибалльной и пятибалльной системе

|  |  |
| --- | --- |
| По десятибалльной шкале | По пятибалльной системе |
| 1 – неудовлетворительно2 – очень плохо3 – плохо | неудовлетворительно – 2 |
| 4 – удовлетворительно5 – весьма удовлетворительно | удовлетворительно – 3 |
| 6 – хорошо7 – очень хорошо | хорошо – 4 |
| 8 – почти отлично9 – отлично10 – блестяще | отлично – 5 |

1. **Учебно-методическое и информационное обеспечение дисциплины**
	1. **Базовый учебник**

отсутствует

* 1. **Основная литература**
* Leach R. Andrew, Molecular Modelling. Priciples and Applications, Willey, 2001,
* Hinchlffe Alan, Modelling Molecular Structures, Willey, 2000
* B. Leimkuhler, Christophe Chipot, Ron Elber, New algorithms for macromolecular simulation,Springer2006
* Wilfred F. van Gunsteren, Paul K. Weiner,Computer simulation of biomolecular systems: theoretical and experimental application, Springer,1997
	1. **Дополнительная литература**

Статья в журнале:

* Funnel metadynamics as accurate binding free-energy method Vittorio Limongelli, Massimiliano Bonomi, Michele Parrinello Proc Natl Acad Sci U S A. 2013 Apr 16; 110(16): 6358–6363. Published online 2013 Apr 3. doi: 10.1073/pnas.1303186110
* Cooper, S.; Khatib, F.; Treuille, A.; Barbero, J.; Lee, J.; Beenen, M.; Leaver-Fay, A.; Baker, D.; Popović, Z.; Players, F. (2010). "Predicting protein structures with a multiplayer online game". Nature. 466 (7307): 756–760. Bibcode:2010Natur.466..756C. doi:10.1038/nature09304. PMC 2956414. PMID 20686574
* Kuhlman, Brian; Dantas, Gautam; Ireton, Gregory C.; Varani, Gabriele; Stoddard, Barry L.; Baker, David (21 November 2003). "Design of a Novel Globular Protein Fold with Atomic-Level Accuracy". Science. 302 (5649): 1364–1368. doi:10.1126/science.1089427. Retrieved 2 August 2018.
* Mol Interv. 2009 Feb;9(1):22-30. doi: 10.1124/mi.9.1.7. Fragment-based ligand discovery. Fischer M1, Hubbard RE.
* Ten-microsecond molecular dynamics simulation of a fast-folding WW domain. Peter L. Freddolino, Feng Liu, Martin Gruebele, and Klaus Schulten. Biophysical Journal, 94:L75-L77, 2008.
* Force field bias in protein folding simulations. Peter L. Freddolino, Sanghyun Park, Benoit Roux, and Klaus Schulten. Biophysical Journal, 96:3772-3780, 2009.
* Common structural transitions in explicit-solvent simulations of villin headpiece folding. Peter L. Freddolino and Klaus Schulten. Biophysical Journal, 97:2338-2347, 2009.
* Going beyond clustering in MD trajectory analysis: an application to villin headpiece folding. Aruna Rajan, Peter L. Freddolino, and Klaus Schulten. PLoS One, 5:e9890, 2010. (12 pages).
* Challenges in protein folding simulations. Peter L. Freddolino, Christopher B. Harrison, Yanxin Liu, and Klaus Schulten. Nature Physics, 6:751-758, 2010.

* 1. **Справочники, словари, энциклопедии**

www.wikipedia.org

* 1. **Программные средства**

Для успешного освоения дисциплины, студент использует следующие программные средства:

* Python 2/3, scipy numpy модули.
* GPL Gromacs, Autodock Vina, Modeller, PyMol
* Orca (Free for academic use)
	1. **Дистанционная поддержка дисциплины**

Сайт с заданиями и лекциями : <https://vsb.fbb.msu.ru/>

Удаленный сервер для выполнения заданий: https://shadbox.vsb.fbb.msu.ru/

1. **Материально-техническое обеспечение дисциплины**

Проектор, компьютерный класс, сервер с Jupyter Hub

Приложение 1

**Примеры подсчета оценки за дисциплину в различных случаях**

Накопленная оценка за текущий контроль учитывает результаты студента по текущему контролю следующим образом:

О*накопленная*= k1\* *Отекущий* + k2\* Оауд + k3\* Осам.работа

где*Отекущий* рассчитывается как взвешенная сумма всех форм текущего контроля, предусмотренных в ОУП

*Отекущий* = *n1·Оэссе + n2·Ок/р + n3·Ореф + n4·Окол + n5·Одз* ;

[Оставьте те формы текущего контроля, которые предусмотрены в ОУП. сумма удельных весов должна быть равна единице: ∑ni = 1] Способ округления накопленной оценки текущего контроля: [указывается способ – арифметический, в пользу студента, другое].

Результирующая оценка за дисциплину рассчитывается следующим образом:

1. Если дисциплина преподается один модуль:

*Орезульт = k1\* Онакопл + k2 \*·Оэкз*

Способ округления накопленной оценки промежуточного (завершающего) контроля: [указывается способ – арифметический, в пользу студента, другое].

1. Если дисциплина преподается несколько модулей (например, 3):

*Опромежуточная i* = *m1·Отекущая i этапа + m2·Опромежуточный экзамен*

Где *Отекущая i этапа* рассчитывается по приведеннойвыше формуле

О*накопленная завершающая =* (О*промежуточная 1+* О*промежуточная 2+* О*накопленная 3):на число модулей*

Где О*промежуточная 1+* О*промежуточная 2 –* промежуточные оценки этапов 1 и 2,
а О*накопленная 3 –* накопленная оценка последнего этапа перед завершающим экзаменом

Способ округления накопленной оценки промежуточного (завершающего) контроля в форме экзамена: [указывается способ – арифметический, в пользу студента, другое].

[Сумма удельных весов должна быть равна единице: ∑mi = 1, при этом, 0,2 ≤ *m1 ≤* 0,8 (согласно Положению об организации промежуточной аттестации и текущего контроля успеваемости студентов НИУ ВШЭ, утвержденному приказом НИУ ВШЭ от 19.08.2014 №6.18.1-01/1908-02**)**

ОПЦИОНАЛЬНО: На экзамене студент может получить дополнительный вопрос (дополнительную практическую задачу, решить к пересдаче домашнее задание), ответ на который оценивается в 1 балл.

[Оставьте те оценки, которые учитываются при выставлении результирующей оценки за промежуточный или завершающий контроль. Сумма удельных весов должна быть равна единице: ∑ki = 1, при этом, **0,2 ≤ *k1 ≤* 0,8 После всех формул в обязательном порядке приводится способ округления полученного результата.**]

[Только для многомодульных дисциплин, по которым предусмотрен промежуточный контроль, укажите один из предложенных вариантов формирования оценки, которая идет в диплом]

1. Указать другие виды аудиторной работы студентов, если они применяются при изучении данной дисциплины. [↑](#footnote-ref-0)