РЕКОМЕНДУЕМАЯ ФОРМА для разработчиков
основных профессиональных образовательных программ
при реализации ОС МГУ на основе ФГОС 3++ и ФГОС 3+

Федеральное государственное бюджетное образовательное учреждение высшего
профессионального образования

Московский государственный университет имени М.В. Ломоносова

факультет, институт, кафедра и т.д.

УТВЕРЖДАЮ

(указать должность)

_____/_____
_ /

« ____ » _____ 20 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Наименование дисциплины (модуля):

**Методы машинного обучения для обработки биологических и медицинских
изображений**

код и наименование дисциплины (модуля)

Уровень высшего образования:

Магистратура

указывается: бакалавриат, магистратура или специалитет

Направление подготовки (специальность):

33.05.01 – фармация

(код и название направления/специальности)

Направленность (профиль) ОПОП:

Регенеративная биомедицина

(если дисциплина(модуль) относится к вариативной части программы)

Форма обучения:

очная

очная, очно-заочная, заочная

Рабочая программа рассмотрена и одобрена

Методической комиссией факультета, на заседании кафедры и т.п.

(протокол № _____, дата)

Москва 20__

На обратной стороне титула:

Рабочая программа дисциплины (модуля) разработана в соответствии с самостоятельно установленным МГУ образовательным стандартом (ОС МГУ) для реализуемых основных профессиональных образовательных программ высшего образования по специальности «Фармация».

ОС МГУ утвержден решением Ученого совета МГУ имени М.В. Ломоносова от _____
20 ____ года (протокол № __).

Год (годы) приема на обучение _____

1. Место дисциплины (модуля) в структуре ОПОП: Вариативная часть ОПОП, дисциплина по выбору.

2. Входные требования для освоения дисциплины (модуля), предварительные условия: базовые знания программирования на языке Python (рекомендуется).

3. Планируемые результаты обучения по дисциплине (модулю), соотнесенные с требуемыми компетенциями выпускников и индикаторами (показателями) их достижения.

Компетенции выпускников (коды)	Планируемые результаты обучения по дисциплине (модулю), соотнесенные с компетенциями
	<i>Знать:</i>
	<i>Знать:</i>

*) Индикаторы (показатели) достижения компетенций переносятся в данную таблицу из документа «Оценочные и методические материалы для контроля формирования компетенций у обучающихся в процессе освоения образовательной программы», входящего в комплект документов, составляющих основную профессиональную образовательную программу по направлению подготовки «_____».

4. Объем дисциплины (модуля) в зачетных единицах с указанием количества академических часов (астрономических часов), выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся:

Объем дисциплины (модуля) составляет 2 зачетные единицы, всего 72 часа, из которых 60 часов составляет контактная работа обучающегося с преподавателем (30 часов занятия лекционного типа и 30 часов семинарского типа), 10 часов составляет самостоятельная работа обучающегося, 2 часа отведено на промежуточную и итоговую аттестацию обучающихся.

5. Формат обучения лекции в виде презентаций

6. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических или астрономических часов и виды учебных занятий

Наименование и краткое содержание раздела в и тем дисциплины (модуля)	Всего (часы)	В том числе							
		Контактная работа (работа во взаимодействии с преподавателем), часы						Самостоятельная работа обучающегося, часы	
		из них						из них	
форма промежуточной аттестации по дисциплине (модулю)		Занятия лекционного типа	Занятия семинарского типа	Групповые консультации	Индивидуальные консультации	Учебные занятия, направленные на проведение текущего контроля успеваемости и промежуточной аттестации)*	Всего	Выполнение домашних заданий, подготовка рефератов и т.п.	Всего
Введение в Python	8	4	4				8		
Традиционные	8	4	4				6	2	2

методы обработк и изображ ений									
Базовые принцип ы машинно го обучения	5	2	2				4	1	1
Традици онный machine learning в обработк е изображ ений	8	4	4				6	2	2
Основы глубоког о машинно го обучения	6	2	2				4	2	2
Нейронн ые сети для классиф икации изображ ений	9	4	4				8	1	1
Принцип ы semantic segmenta	4	2	2				4		

tion									
Принципы Object detection	5	2	2				4	1	1
Принципы Instance segmentation	5	2	2				4	1	1
Принципы работы Pix2pix моделей	4	2	2				4		
Autoencoders	4	2	2				4		
Итоговая аттестация	2						2		
Итого	72	30	30				62	10	10

**Текущий контроль успеваемости может быть реализован в рамках занятий семинарского типа, групповых или индивидуальных консультаций*

*** Промежуточная аттестация может проходить как в традиционных форма (зачет, экзамен), так и в иных формах (балльно-рейтинговая система, защита портфолио, комплексный экзамен, включающий выполнение практических заданий (возможно наряду с традиционными ответами на вопросы по программе дисциплины (модуля)))*

Применение методов машинного обучения для обработки биологических и медицинских изображений

(1 лекция) Введение в Python. Знакомство с синтаксисом (if, for, оформление функций), типами данных (mutable, immutable), объектно ориентированным программированием (оформление классов, наследование, магические методы).

Практическое занятие: обработка работы с классами и типами данных.

(1 лекция) Введение в Python. Работа с библиотеками numpy, pandas (работа с размерностями). Визуализация данных с помощью plt.plot и seaborn.

Практическое занятие: создание pandas датасетов, взаимодействие между numpy и pandas, анализ табличных данных.

(2 лекции) Традиционные методы обработки изображений. Что такое изображение? Типы изображений, форматы хранения (float, int8, int32) и обработки (PIL.Image, numpy.array, torch.tensor). Threshold, понятие о бинарном слое. Базовые операции с изображениями. Стандартные библиотеки python для работы с изображениями (Open CV, torchvision.transforms, PIL .Image). Аугментация.

Практическое занятие: работа с изображениями, преобразование форматов, threshold, бинарные слои.

(1 лекция) Традиционные методы обработки изображений. Понятие о feature extraction. Продвинутое операции с бинарными слоями. Понятие о свертке. Фильтр Гаусса, Edge detection, entropy-based thresholding.

Практическое занятие: применение нескольких традиционных методов обработки изображений к реальной задаче.

(2 лекции) Традиционный machine learning в обработке изображений. Типы ML-задач: классификация, кластеризация, регрессия, сегментация и детекция. Разделение данных на train, validation, test. k-fold cross-validation. Понятие о гиперпараметрах модели. Метрики (accuracy, precision, recall, ROC AUC, MSE, MAE). Дисбаланс классов. K-nearest neighbors. Нормализация.

Практическое занятие: применение алгоритмов K-nearest neighbors для классификации изображений.

(1 лекция) Традиционный machine learning в обработке изображений. Ансамблевые модели на основе деревьев: Random forest, градиентный boosting. Optuna.

Практическое занятие: работа с Random Forest Classifier на реальных данных.

(1 лекция) Базовые принципы машинного обучения на примере линейной регрессии. Основы библиотеки Torch. Loss - функция. Градиентный спуск. Оптимизаторы SGR и ADAM. Переобучение, L1, L2 - регуляризация. Support Vector Machines. Наивный Байесовский классификатор. Кросс-энтропия. Взрыв и затухание градиента.

Практическое занятие: обучение модели линейной регрессии на реальных данных.

(1 лекция) Основы глубокого машинного обучения. Перцептрон. Функции активации. Теорема об универсальном аппроксиматоре. Полносвязные нейронные сети.

Практическое занятие: создание и обучение маленькой нейронной сети. Оценка влияния ошибок в предобработке данных на качество обучения (разбор случая при обучении на ненормализованных данных, при сильном дисбалансе классов).

(2 лекции) Нейронные сети для классификации изображений. Сверточные нейронные сети. Слои по-разному ведущие себя в режиме train и eval: Dropout (1d и 2d), Batch Normalization (1d и 2d). Residual блок.

Практическое занятие: написание нейросети для классификации изображений на примере архитектуры ResNet.

(1 лекция) Нейронные сети для классификации изображений. Современные модели в классификации изображений (torch image model: EfficientNet, VT), принципы их работы и обучения под свои задачи. Fine Tuning.

Практическое занятие: знакомство с работой моделей, загрузка и дообучение предобученной модели, сохранение весов.

(1 лекция) Различные типы сегментации. Принципы semantic segmentation. Segmentation model pytorch. Метрики качества для сегментации (torch.metrics).

Практическое занятие: создание и обучение модели U-Net на реальных данных.

(1 лекция) Принципы Object detection. Метрики качества для Object detection.

Практическое занятие: дообучение и применение модели YOLOv6v3 на реальных данных.

(1 лекция) Принципы Instance segmentation. Cellpose, Stardist. Mask-RCNN. Panoptic segmentation.

Практическое занятие: дообучение и применение модели из Cellpose model zoo на реальных данных.

(1 лекция) Принципы работы Pix2pix моделей.

Практическое занятие: обучение и применение Pix2pix модели на реальных данных.

(1 лекция) Supervised and unsupervised learning. Autoencoders. Традиционные методы денойзинга.

Практическое занятие: обработка изображений с высоким уровнем шума.

(1 лекция) Explainability. Оценка важности признаков для простых моделей. SHAP и градиентные методы.

7. Фонд оценочных средств для оценивания полученных студентом результатов обучения по дисциплине (модулю)

Компетенции и выпускников (коды)	Планируемые результаты обучения по дисциплине (модулю), соотнесенные с компетенциями	Виды оценочных средств	Шкалы оценивания			
			2	3	4	5
	<i>Знать:</i>	Устный опрос, Зачет	Не знает	Знает фрагментарно	Знает, но допускает неточности	Полностью знает
	<i>Знать:</i>	Устный опрос, Зачет	Не знает	Знает фрагментарно	Знает, но допускает неточности	Полностью знает

					ости	
	<i>Знать:</i>	Устный опрос, Зачет	Не знае т	Знает фрагмен тарно	Знает, но допуск ает неточн ости	Полно стью знает
	<i>Знать:</i>	Устный опрос, Зачет	Не знае т	Знает фрагмен тарно	Знает, но допуск ает неточн ости	Полно стью знает
	<i>Знать:</i>	Устный опрос, Зачет	Не знае т	Знает фрагмен тарно	Знает, но допуск ает неточн ости	Полно стью знает

ПРИМЕРЫ ВОПРОСОВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

1. Методы чтения изображения в Python
2. Методы итерации по файлам в папке в Python
3. Расширения цифровых изображений: .jpeg, .png, .tiff
4. Понятие бинарного слоя
5. Логические операции с бинарными слоями
6. Принципы train-val-test split
7. Принципы k-fold cross-validation

8. Принцип градиентного спуска
9. Отличия supervised и unsupervised learning
10. Основные ML-задач. К каким основным типам задач сводится сегментация и детекция?
11. Причины и методы борьбы с переобучением.
12. Предобработка данных: на что важно обратить внимание
13. Принципы feature extraction для обработки изображений методами традиционного машинного обучения
14. Принцип работы свёрточных нейронных сетей.
15. Обучение нейронных сетей, алгоритм обратного распространения ошибки.
16. Ускорение сходимости. Оптимизатор Adam.
17. Современные модели классификации изображений. Область применения.
18. Отличия semantic segmentation от instance segmentation, другие типы сегментации.
19. Принципы object detection, современные модели. Область применения.
20. Принципы instance segmentation, фреймворк Cellpose. Область применения.
21. Принципы работы pix2pix моделей, область применения
22. Принцип работы и создания автоэнкодеров, область применения.

8. Ресурсное обеспечение:

1. Russell S. J. Artificial intelligence a modern approach. – Pearson Education, Inc., 2010. Krauss G., Biochemistry of Signal Transduction and Regulation, 2001
2. Hemanth D. J., Estrela V. V. (ed.). Deep learning for image processing applications. – IOS Press, 2017. – Т. 31.
3. Laganiere R. OpenCV computer vision application programming cookbook second edition. – Packt Publishing Ltd, 2014.
4. Dougherty E. R. Digital image processing methods. – CRC Press, 2020.

Интернет-ресурсы.

- “Поколение Python”: курс для начинающих (<https://stepik.org/58852>)
- “Поколение Python”: курс для продвинутых (<https://stepik.org/68343>)

МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ:

А. Помещения: аудитории, оснащённые компьютерами и мультимедийными проекторами.

Б. Оборудование: компьютеры, мультимедийные проекторы.

В. Иные материалы: не требуются.

9. Язык преподавания.

Английский

10. Преподаватели: аспирант Чечехина Е.С., аспирант Волошин Н.С.

11. Разработчики программы: аспирант Чечехина Е.С., аспирант Волошин Н.С.