

**МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
КЫРГЫЗСКОЙ РЕСПУБЛИКИ**

**КЫРГЫЗСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
СТРОИТЕЛЬСТВА, ТРАНСПОРТА И АРХИТЕКТУРЫ им. Н. Исанова**

Кафедра «Прикладная информатика»

**СИЛЛАБУС (Рабочая программа)
дисциплины «Методы машинного обучения»**

для направления: 710100- Компьютерные и информационные технологии
по образовательной программе:
“PhD - Информационные системы и процессы”

Форма обучения: очная

Докторантура	
Семестр	2
Количество кредитов	10
Лекции, часов	30
Лабораторные занятия, часов	60
Самостоятельная работа, часов	210
СРСР	
Формы контроля	экзамен

Бишкек -2022

Рабочая программа предназначена для преподавания дисциплины “Методы машинного обучения” докторантам очной формы обучения по направлению 710100 – Компьютерные и информационные технологии (ОП:Информационные системы и процессы) во 2 семестре.

Составитель. к.ф.-м.н., доцент Орозобекова А.К.

Тематический план

модульной программы по дисциплине
«Методы машинного обучения» для PhD по направлению
710100 «Компьютерные и информационные системы»

№ дисциплинарного модуля	№ темы лекций (лаб)	Лекционные занятия	ЛК	Лабораторные занятия	ЛБ	СРС
		Наименование темы		Наименование темы		
ДМ 1	1	Введение в Машинное обучение	2	Работа с табличными данными и Pandas	2	
	2	Линейная регрессия, функции потерь в регрессии.	2	sklearn и линейная регрессия	4	14
	3	Обобщающая способность, градиентные методы обучения.	2	Матрично-векторное дифференцирование	4	14
	4	Градиентные методы, регуляризация	2	Работа с данными и признаками	4	14
	5	Линейная классификация	2	Метрики классификации	4	16
	6	Метрики качества классификации, логистическая регрессия.	2	Калибровка вероятностей.	4	14
	7	Полиномиальная регрессия. Аппроксимация полиномами высокой степени	2	Решающие деревья	4	14
	8	Многоклассовая классификация	2	Разложение ошибки на смещение и разброс	4	14
ДМ 2	9	Решающие деревья	2	Градиентный бустинг	4	16
	10	Бэггинг, случайные леса и разложение ошибки на смещение и разброс.	2	Градиентный бустинг	6	16
	11	Градиентный бустинг	2	Виды градиентного бустинга: XGB, LightGBM, CatBoost	4	14
	12	Градиентный бустинг. Альтернативный подход.	2	Виды градиентного бустинга: XGB, LightGBM, CatBoost	4	16

	13	Линейные композиции, бустинг и случайные леса	2	Кластеризация	4	16
	14	Градиентный бустинг над деревьями решений	2	Методы понижения размерности: PCE, tSNE	4	16
	15	Кластеризация. Обучение без учителя	2	Методы понижения размерности: PCE, tSNE	4	16
		Всего:	30		60	210

Тематический план составила

к.ф.-м.н., доцент Орозобекова А.К. _

ЦЕЛИ И ЗАДАЧИ ДИСЦИПЛИНЫ

Цель курса (дисциплины): изучение методов машинного обучения, обучения и построение моделей для различных прикладных задач с применением библиотек Python.

Основные задачи курса:

- получение практических навыков по реализации и функционированию машинного обучения;
- изучение ряда подходов построения линейных моделей задач прикладного характера;
- изучение подходов к построению обобщенных линейных моделей с применением полиномов очень высоких степеней
- обучающие модели для линейных процессов, обученная на полиномиальных признаках такие модели
- программная реализация конкретных задач прикладного характера.

Результаты обучения дисциплины: «Методы машинного обучения»

РОД1 – знать основы машинного обучения;

РОД2 – знать теоретические основы регрессионного анализа;

РОД3- владеть основами программирования Python и его библиотеками;

РОД4- знать основы построения линейных моделей задач прикладного характера

В результате изучения дисциплины PhD студенты должны:

знать:

- основы машинного обучения;
- основы построения линейных моделей задач прикладного характера;
- основы программирования на Python и его библиотеки

уметь:

- применить современные технологии построения моделей, регуляризация переобученных моделей, которые могут быть использованы для включения параметров регуляризации.

владеть:

- обучающими моделями для линейных процессов на полиномиальных признаках.
- программной реализацией конкретных задач прикладного характера.

ПРЕДЛАГАЕМЫЕ РЕЗУЛЬТАТЫ ИЗУЧЕНИЯ ДИСЦИПЛИНЫ

- владением методологией теоретических и экспериментальных исследований в области профессиональной деятельности (ОК-3);
- готовность участвовать в работе республиканских и международных исследовательских коллективов по решению научно-исследовательских задач (ИК-2);
- способность к формальной постановке задач системного анализа, оптимизации, управления, принятия решений и обработки информации (ПК-1);
- способностью исследовать прикладные информационные системы для различных областей применения, строить аналитические, процедурные, информационные модели предметной области (ПК-5);

Основная литература

1. Рашка, С. Python и машинное обучение // Себастьян Рашка. – ДМК, 2017. – 418 с.
2. Вьюгин, В.В. Математические основы машинного обучения и прогнозирования // В.В. Вьюгин. – МЦМНО, 2014. - 305 с.
3. Бринк, Х., Ричардс, Дж., Феверолф, М. Машинное обучение // Хенрик Бринк, Джозеф Ричардс, Марк Феверолф. – Питер, 2017. – 336 с.
4. Флах П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных / пер. с англ. А. А. Слинкина. – М.: ДМК Пресс, 2015. – 400 с.: ил.
5. Бурков, А. Машинное обучение без лишних слов // Андрей Бурков. – СПб.: Питер, 2020. – 192 с. Коэльо Л.П., Ричарт В. Построение систем машинного обучения на языке Python. 2016. 302 с.
6. Мерков А. Б. Распознавание образов. Введение в методы статистического обучения. 2011. 256 с.
7. Мерков А. Б. Распознавание образов. Построение и обучение вероятностных моделей. 2014. 238 с.
8. Hastie T., Tibshirani R., Friedman J. The Elements of Statistical Learning. Springer, 2014. — 739 p.
9. Bishop C. M. Pattern Recognition and Machine Learning. — Springer, 2016. — 738 p.

Дополнительная литература:

1. Шарден, Б., Массаро, Л., Боскетти, А. Крупномасштабное машинное обучение вместе с Python // Б. Шарден, Л. Массарон, А. Боскетти. – ДМК Пресс, 2017. – 358 с.
2. Жерон Оуэн Прикладное машинное обучение с помощью Scikit-Learn и TensorFlow: концепции, инструменты и техники для создания интеллектуальных систем. Пер. с англ. - СПб.: ООО "Альфа-книга": 2018. - 688 с.
3. man, L. Arcing classifiers (with discussion), Annals of Statistics 26: 2008. 801–849.

Интернет ресурсы:

1. <https://ya-r.ru/2020/05/07/vorontsov-kurs-mashinnoe-obuchenie-2019-shkola-analiza-dannyh/>
2. <https://www.youtube.com/watch?v=QlktmPA8nb0&list=PLJOzdkh8T5krxc4HsHbB8g8f0hu7973fK&index=3>
3. [http://www.machinelearning.ru/wiki/index.php?title=%D0%9C%D0%B0%D1%88%D0%B8%D0%BD%D0%BD%D0%BE%D0%B5_%D0%BE%D0%B1%D1%83%D1%87%D0%B5%D0%BD%D0%B8%D0%B5_\(%D0%BA%D1%83%D1%80%D1%81_%D0%BB%D0%B5%D0%BA%D1%86%D0%B8%D0%B9,%D0%9A.%D0%92.%D0%92%D0%BE%D1%80%D0%BE%D0%BD%D1%86%D0%BE%D0%B2\)](http://www.machinelearning.ru/wiki/index.php?title=%D0%9C%D0%B0%D1%88%D0%B8%D0%BD%D0%BD%D0%BE%D0%B5_%D0%BE%D0%B1%D1%83%D1%87%D0%B5%D0%BD%D0%B8%D0%B5_(%D0%BA%D1%83%D1%80%D1%81_%D0%BB%D0%B5%D0%BA%D1%86%D0%B8%D0%B9,%D0%9A.%D0%92.%D0%92%D0%BE%D1%80%D0%BE%D0%BD%D1%86%D0%BE%D0%B2))

МЕТОДЫ ОБУЧЕНИЯ ДИСЦИПЛИНЫ

На кафедре при преподавании дисциплины применяются следующие методы обучения:

- устное изложение учебного материала на лекциях;
- самостоятельное изучение студентами учебного материала по рекомендованной литературе;
- выполнение контрольных работ PhD студентами.

Выбор методов проведения занятий обусловлен учебными целями, содержанием учебного материала, временем, отводимым на занятия.

На занятиях в тесном сочетании применяется несколько методов, один из которых выступает ведущим. Он определяет построение и вид занятий.

На лекциях излагаются лишь основные, имеющие принципиальное значение и наиболее трудные для понимания и усвоения теоретические и расчетно-конструкторские вопросы.

Теоретические знания, полученные PhD студентами на лекциях и при самостоятельном изучении курса по литературным источникам, закрепляются при выполнении индивидуальных контрольных работ.

При выполнении индивидуальных контрольных работ обращается особое внимание на выработку у PhD студентов умения пользоваться нормативной и справочной литературой, грамотно выполнять и оформлять инженерные расчеты и чертежи и умения обрабатывать отчетные документы в срок и с высоким качеством.

СРЕДСТВА ОБУЧЕНИЯ

К средствам обучения по данной дисциплине относятся:

- речь преподавателя;
- технические средства обучения: проектор (доска), персональные компьютеры;
- учебники, учебные пособия, справочники, изданные лекции, методические указания;

На занятиях по дисциплине должны широко использоваться разнообразные средства обучения, способствующие более полному и правильному пониманию темы лекции или лабораторного занятия.