

ВВЕДЕНИЕ В МАШИННОЕ ОБУЧЕНИЕ

Преподаватели: **Бурнаев Евгений Владимирович,**

Зайцев Алексей Алексеевич

1. Место и статус дисциплины в структуре образовательной программы

Уровень обучения: *3-4/Специалитет, 3-4/Бакалавриат*

Язык обучения: *Русский*

Формат обучения: *онлайн*

Время для консультаций: *по договоренности с преподавателем*

2. Входные требования для освоения дисциплины

- *Линейная алгебра*
- *Математический анализ*
- *Теория вероятности*
- *Методы оптимизации*
- *Программирование (Python или знание другого скриптового языка)*

3. Аннотация дисциплины

Курс представляет собой общее введение в машинное обучение (МО) и его приложения. Курс охватывает фундаментальные темы в МО и описывает наиболее важные алгоритмические основы и подходы, аспекты применения алгоритмов. Курс начинается с обзора канонических приложений и задач МО, сценариев обучения и т.д. Далее, в курсе подробно обсуждаются фундаментальные алгоритмы МО для классификации, регрессии, кластеризации и т.д., их свойства и практическое применение. Последняя часть курса посвящена продвинутым темам МО, таким как гауссовские процессы, нейронные сети. В практических сессиях мы покажем, как использовать методы МО и настраивать их гиперпараметры. Домашние задания включают в себя применение существующих алгоритмов для решения задач анализа данных. Предполагается, что студенты, посещающие данный курс, уже знакомы с основными понятиями линейной алгебры, теории вероятностей, математического анализа, оптимизации и программирования на python.

4. Планируемые результаты обучения по дисциплине

- *Иметь хорошее понимание фундаментальных проблем и проблем МО: анализ свойств данных, выбор типа модели и сложности соответствующего функционального класса, и т.п.;*
- *Иметь представление о сильных и слабых сторонах многих популярных подходов МО;*
- *Понимать лежащие в основе МО математические концепции, позволяющие проводить обучение с учителем и обучение без учителя;*

- Уметь разрабатывать и использовать алгоритмы машинного обучения для решения прикладных задач.

5. Содержание дисциплины

№	Темы	Кол-во занятий	
		Лекции	Семинары/ практикумы
1	Введение	1	1
2	Регрессия, Ядерный метод	1	1
3	Классификация	1	1
4	Метод опорных векторов	1	1
5	Дерево решений. Случайный лес	1	1
6	Несбалансированная и многоклассовая классификация	1	1
7	Бустинг	2	2
8	Выбор модели и признаков. Снижение размерности	2	2
9	Введение в нейронные сети	1	1
10	Глубокие нейронные сети	1	1
11	Выявление аномалий	1	1
12	Кластеризация	1	1
	ИТОГО	14	14

6. Элементы контроля и формула расчета финальной оценки:

- Посещаемость
- Домашняя работа (3 задания)
- Финальный проект

$$\text{Оценка} = 0.05 * \text{Посещаемость} + 0.55 * \text{Домашняя работа} + 0.4 * \text{Финальный проект}$$

8-10	75-100%	5	отл
6-7	65-75%	4	хорошо
4-5	55-65%	3	удовл.
0-3	<55%	0-2	неудовл.

7. Список литературы

а. Рекомендуемая основная литература

- Trevor Hastie, Robert Tibshirani, Jerome Friedman. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer, 2009.
- Shai Shalev-Shwartz, Shai Ben-David. Understanding Machine Learning: From Theory to Algorithms. Cambridge, 2014.
- R.E. Schapire, Y. Freund. Boosting. MIT, 2012
- M. Mohri, A. Rostamizadeh, A. Talwalkar. Foundations of Machine Learning. MIT, 2012.
- B. Clarke, E. Fokoue, H.H. Zhang. Principles and Theory for Data Mining and Machine Learning. Springer, 2009
- Kevin P. Murphy. Machine Learning: A Probabilistic Perspective. MIT Press, 2012.
- Sutton and Barto. Reinforcement Learning: An Introduction. MIT Press, 1998.

б. Рекомендуемая дополнительная литература

- <http://web4.cs.ucl.ac.uk/staff/D.Barber/pmwiki/pmwiki.php?%20n=Brml.HomePage>
- <http://gaussianprocess.org/gpml/>
- <http://wol.ra.phy.cam.ac.uk/mackay/itila/book.html>

8. Описание материально-технической базы

Ноутбук с установленным Google Colab

9. Примеры заданий текущего контроля

<https://ucarecdn.com/a3ac299d-03e6-40c0-b070-54ebd213eed1/>