

Алгоритмы машинного обучения

Программа курса

Лектор: к.ф.-м.н., доцент Китов Виктор Владимирович.

Инструктор: Виктор Антипов

Цель курса: изучение основных алгоритмов в анализе данных, прогнозировании и машинном обучении, необходимых для чтения литературы и построения собственных интеллектуальных систем в финансовой области.

Пререквизиты: базовое знание математического анализа, линейной алгебры, теории вероятностей и математической статистики. Понадобятся знакомство с понятиями: производная, интеграл, вектор, матрица, операции над ними, собственные значения и собственные вектора, определитель, ранг, обратная матрица, основные вероятностные распределения, метод максимального правдоподобия.

Краткое содержание курса

1. Основные понятия и применения машинного обучения.
2. Метрические методы. Возможности оптимизации.
3. Сложность моделей. Подготовка данных и генерация признаков.
4. Линейная регрессия и её обобщения. Метод стохастического градиентного спуска.
5. Линейная классификация.
6. Решающие деревья. Оценка качества классификаторов.
7. Ансамбли прогнозирующих алгоритмов
8. Алгоритм бустинга и его модификации.
9. Нейронные сети — архитектура и оптимизация.
10. Техники настройки глубоких нейронных сетей. Генеративно-сопоставительные сети.
11. Метод главных компонент. Сингулярное разложение. Нелинейные методы снижения размерности.
12. Кластеризация. Детекция аномалий.

Оценивание

Баллы за курс выставляются по формуле:

$score = 0.4 * homework + 0.3 * midterm + 0.3 * exam$, где:

homework - баллы за домашние задания

midterm - баллы за промежуточный экзамен

exam - баллы за письменный экзамен

Оценка ставится по критерию:

5: $score \geq 75\%$

4: $60\% \leq score < 75\%$

3: $40\% \leq score < 60\%$

Промежуточный и итоговый экзамен пишутся письменно, разрешено использование лекционных материалов, учебников, собственных конспектов. Общение и использование интернета не разрешено.

Литература

- Trevor Hastie, Robert Tibshirani, Jerome Friedman. The Elements of Statistical Learning. 2nd edition. Springer. 2009. <https://web.stanford.edu/~hastie/ElemStatLearn>.
- Charu C. Aggarwal. Data Mining: The Textbook. Springer. 2015.
- Ian Goodfellow, Yoshua Bengio, Aaron Courville. Deep Learning. MIT Press. 2016. <https://www.deeplearningbook.org>