

КАЗАХСКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ ИМЕНИ АЛЬ-ФАРАБИ

А.Т. Агишев

МЕТОДЫ И ТЕХНОЛОГИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

**Сборник лекций для студентов магистратуры,
обучающихся по образовательной программе
«7M06201 - Радиотехника, электроника и телекоммуникации»**

Алматы, 2025

Лекция 6. Концепции машинного обучения

Цель лекции

Сформировать у студентов системное понимание ключевых концепций машинного обучения, лежащих в основе построения, обучения и оценки моделей. Рассмотреть понятия обучающей выборки, обобщающей способности, функции потерь, переобучения и смещения, а также связи между bias–variance tradeoff и эффективностью моделей.

Основные вопросы:

1. Основные понятия и этапы машинного обучения.
2. Понятие обучающей, валидационной и тестовой выборки.
3. Обобщающая способность модели (generalization).
4. Ошибки обучения: смещение (bias) и разброс (variance).
5. Проблема переобучения (overfitting) и недообучения (underfitting).
6. Функции потерь и оптимизация.
7. Регуляризация и методы предотвращения переобучения.
8. Баланс между точностью и сложностью модели.

Краткие тезисы:

1. **Основные понятия.**
По Alpaydin (2020) и Russell & Norvig (2021), процесс машинного обучения включает три ключевых элемента:
 - **Представление (Representation):** способ описания данных и модели;
 - **Оценка (Evaluation):** метрика качества (loss, accuracy и др.);
 - **Оптимизация (Optimization):** метод поиска параметров, минимизирующих ошибку.Эти три компонента образуют «треугольник машинного обучения».
2. **Обучающая, валидационная и тестовая выборки.**
Модель обучается на *training set*, настраивается по *validation set* и оценивается по *test set*.
 - **Training set** — данные, используемые для подбора параметров модели.
 - **Validation set** — проверка модели во время обучения для настройки гиперпараметров.
 - **Test set** — независимая проверка качества после завершения обучения.Типичное соотношение: 70/15/15 или 80/10/10. (Goodfellow et al., 2016, гл. 5)
3. **Обобщающая способность (Generalization).**
Обобщение — это способность модели давать правильные прогнозы на

новых, ранее не встречавшихся данных.
Качество обобщения зависит от:

- сложности модели;
- объёма и разнообразия обучающей выборки;
- регуляризации и шумоустойчивости данных.

Модель с высокой обобщающей способностью не запоминает данные, а выявляет закономерности.

4. Смещение и разброс (Bias–Variance Tradeoff).

- **Bias (смещение)** — ошибка из-за чрезмерных упрощений модели (underfitting).
- **Variance (разброс)** — чувствительность модели к шуму и случайным флуктуациям (overfitting).

Суммарная ошибка модели складывается из трёх компонентов:

$$E_{total} = \text{Bias}^2 + \text{Variance} + \text{Irreducible Error}$$

Цель обучения — найти оптимальный баланс между bias и variance.

Пример: простая линейная регрессия имеет высокий bias и низкий variance; глубокая нейросеть — наоборот.

5. Проблема переобучения и недообучения.

- **Underfitting:** модель слишком проста, не улавливает закономерностей.
- **Overfitting:** модель слишком сложна и подстраивается под шумы обучающей выборки.
Признаки overfitting — высокая точность на тренировке, но низкая на тесте.

Методы борьбы:

- регуляризация (L1/L2),
- dropout,
- увеличение данных (data augmentation),
- ранняя остановка (*early stopping*),
- кросс-валидация.

6. Функции потерь (Loss Functions).

Функция потерь измеряет расхождение между предсказанием и истиной.

Примеры:

- **MSE (Mean Squared Error)** — для регрессии:

$$L = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

- **Cross-Entropy Loss** — для классификации:

$$L = - \sum_{i=1}^n y_i \log (\hat{y}_i)$$

7. Оптимизация проводится методами градиентного спуска или его модификаций (SGD, Adam, RMSProp).

7. **Регуляризация и предотвращение переобучения.**
Регуляризация добавляет штраф за сложность модели:

$$L_{reg} = L + \lambda \| \theta \|^2$$

где λ — коэффициент регуляризации.

- **L1 (Lasso)** — обнуляет малозначимые параметры (спарсность).
- **L2 (Ridge)** — сглаживает веса, предотвращая скачки.
- **Dropout** — случайное выключение нейронов при обучении (Goodfellow et al., гл. 7).

Эти методы повышают устойчивость и способность к обобщению.

8. **Баланс между точностью и сложностью модели.**
По принципу Оккама, предпочтительна модель с минимальной сложностью при приемлемой точности. Сложные модели (глубокие сети) требуют больших данных и регуляризации, простые — могут недообучаться. Инженерная практика требует выбора модели по принципу:

«Достаточно точна, чтобы решать задачу, но не избыточна по сложности.»

Вопросы для контроля, изучаемого материал:

- 1) Что представляют собой три ключевые составляющие машинного обучения: представление, оценка и оптимизация?
- 2) Какова роль обучающей, валидационной и тестовой выборок?
- 3) Что означает понятие “обобщающая способность модели”?
- 4) Какова взаимосвязь между bias и variance?
- 5) Чем отличаются переобучение и недообучение?
- 6) Какие функции потерь применяются для классификации и регрессии?
- 7) Что такое регуляризация и какие методы применяются для её реализации?
- 8) Почему важно учитывать баланс между точностью и сложностью модели?

Рекомендуемый список литературных источников:

1. Russell, S., Norvig, P. Artificial Intelligence: A Modern Approach. 4th Edition. Pearson, 2021.
2. Goodfellow, I., Bengio, Y., Courville, A. Deep Learning. MIT Press, 2016.
3. Alpaydin, E. Introduction to Machine Learning. 4th Edition. MIT Press, 2020.