За 5-10 в лк получим билеты и случайно номер получим

8 30 к конфе и идентификация: на камеру документ

Письменно на вопросы билета отвечаем, скан ответов и отправить в лк 1 файл в пдф

Ответы письменно при включенной камере

Пожелание: разборчивый почерк

До 10 20 в лк получить ответы

2 теории и задача, но ее делать не надо

Не списывать тупо, желательно свои мысли, с пониманием

Заканчиваем писать, в чат сообщить что закончили и занимаемся сканированием, можно камеру на это время отключить

Дедалйн проверки +- 13 00, в лк оценка

До 13 30 в лк сообщить, что не устраивает оценка, тогда лично в зум, там теория и практические вещи (библиотеки, модели и тд)

1. **Понятие машинного обучения.**

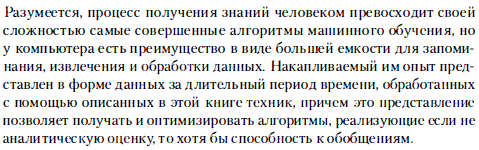
Машинное обучение – одна из форм искусственного интеллекта, включающего робототехнику, обработку лингвистической информации и системы машинного зрения. Машинное обучение относится к специализированной совокупности знаний и связанным с ней техником. Определение Тома Митчела: компьютерная программа обучается, если ее производительность при выполнении определенной задачи, выражаемая в измеряемых единицах, увеличивается по мере накопления опыта.

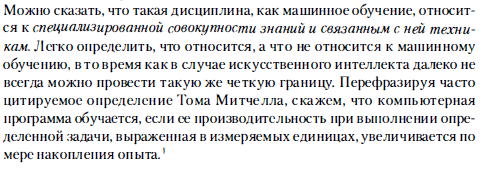
**Машинное обучение** ([англ.](https://ru.wikipedia.org/wiki/%D0%90%D0%BD%D0%B3%D0%BB%D0%B8%D0%B9%D1%81%D0%BA%D0%B8%D0%B9_%D1%8F%D0%B7%D1%8B%D0%BA) *machine learning*, ML) — класс методов [искусственного интеллекта](https://ru.wikipedia.org/wiki/%D0%98%D1%81%D0%BA%D1%83%D1%81%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D1%8B%D0%B9_%D0%B8%D0%BD%D1%82%D0%B5%D0%BB%D0%BB%D0%B5%D0%BA%D1%82), характерной чертой которых является не прямое решение задачи, а обучение в процессе применения решений множества сходных задач. Для построения таких методов используются средства [математической статистики](https://ru.wikipedia.org/wiki/%D0%9C%D0%B0%D1%82%D0%B5%D0%BC%D0%B0%D1%82%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%B0%D1%8F_%D1%81%D1%82%D0%B0%D1%82%D0%B8%D1%81%D1%82%D0%B8%D0%BA%D0%B0), [численных методов](https://ru.wikipedia.org/wiki/%D0%A7%D0%B8%D1%81%D0%BB%D0%B5%D0%BD%D0%BD%D1%8B%D0%B5_%D0%BC%D0%B5%D1%82%D0%BE%D0%B4%D1%8B), [математического анализа](https://ru.wikipedia.org/wiki/%D0%9C%D0%B0%D1%82%D0%B5%D0%BC%D0%B0%D1%82%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%B8%D0%B9_%D0%B0%D0%BD%D0%B0%D0%BB%D0%B8%D0%B7), [методов оптимизации](https://ru.wikipedia.org/wiki/%D0%9E%D0%BF%D1%82%D0%B8%D0%BC%D0%B8%D0%B7%D0%B0%D1%86%D0%B8%D1%8F_(%D0%BC%D0%B0%D1%82%D0%B5%D0%BC%D0%B0%D1%82%D0%B8%D0%BA%D0%B0)), [теории вероятностей](https://ru.wikipedia.org/wiki/%D0%A2%D0%B5%D0%BE%D1%80%D0%B8%D1%8F_%D0%B2%D0%B5%D1%80%D0%BE%D1%8F%D1%82%D0%BD%D0%BE%D1%81%D1%82%D0%B5%D0%B9), [теории графов](https://ru.wikipedia.org/wiki/%D0%A2%D0%B5%D0%BE%D1%80%D0%B8%D1%8F_%D0%B3%D1%80%D0%B0%D1%84%D0%BE%D0%B2), различные техники работы с [данными в цифровой форме](https://ru.wikipedia.org/wiki/%D0%94%D0%B0%D0%BD%D0%BD%D1%8B%D0%B5_(%D0%B2%D1%8B%D1%87%D0%B8%D1%81%D0%BB%D0%B8%D1%82%D0%B5%D0%BB%D1%8C%D0%BD%D0%B0%D1%8F_%D1%82%D0%B5%D1%85%D0%BD%D0%B8%D0%BA%D0%B0)).

Для обучения и обобщения, люди используют различные данные из примеров, включая формы, цвета, текстуры, пропорции и другие характеристики. МО также применяет множество стратегий в различных комбинациях в зависимости от поставленной задачи. Эти стратегии нашли воплощение в наборе алгоритмов, разработанных в самых разных дисциплинах: от статистики, робототехники и прикладной математики до поиска в Интернет, развлекательной сферы, цифровой рекламы и переводов с 1 языка на другой.

Алгоритмы крайне разнообразны, имеют сильные и слабые стороны. Некоторые относят объекты к определенному классу, другие предсказывают числовые значения. Существуют алгоритмы, определяющие сходства и различия допускающих сравнения сущностей. При этом, все алгоритмы обучаются на примерах (т.е. опыте), т.е. способны к обобщению.

Машинное обучение применяется к широкому кругу экономических задач - от обнаружения мошенничества до выбора целевой аудитории и рекомендаций товара, наблюдения за производством в реальном времени, анализа тональности текстов и медицинской диагностики. Оно может взять на себя задачи, которые невозможно выполнить вручную из-за огромного количества подлежащих обработке данных. В случае больших наборов данных машинное обучение иногда обнаруживает не очевидные зависимости, которые невозможно распознать при сколь угодно скрупулезном ручном рассмотрении. При этом комбинация множества таких «слабых» соотношений дает прекрасно работающие механизмы прогнозирования.

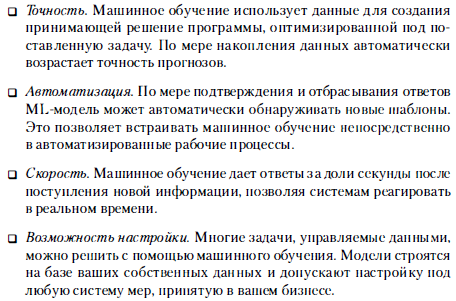




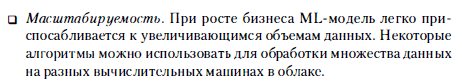
1. **Задачи, решаемые с применением машинного обучения.**

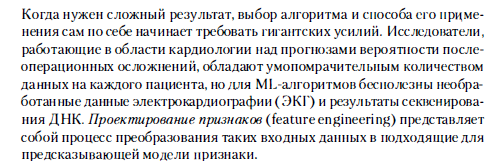


1. **Преимущества машинного обучения. Сложности, возникающие в процессе машинного обучения.**

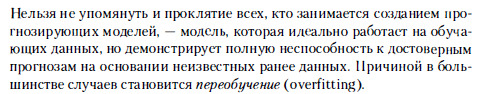
Преимущества

Сложности

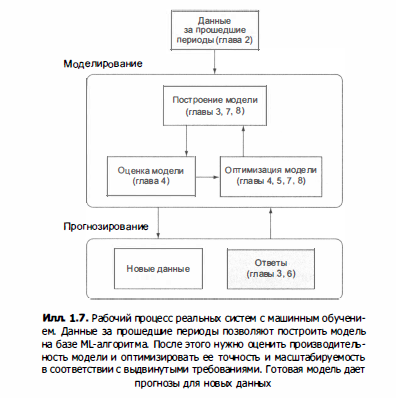
* получить данные в годной к употреблению форме сформулировать проблему таким образом, чтобы к ней можно было применить методы машинного обучения и получить имеющие практическую ценность и измеримые результаты.
* Выбор алгоритма и способ его применения для получения сложного результата



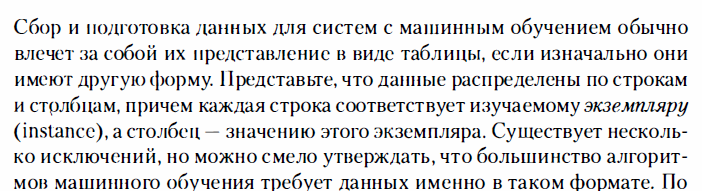
* Переобучение

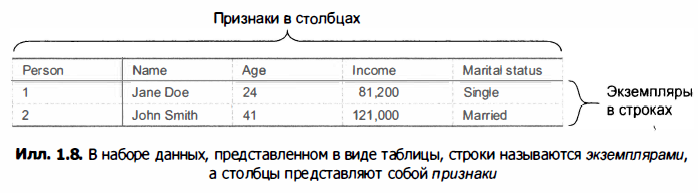


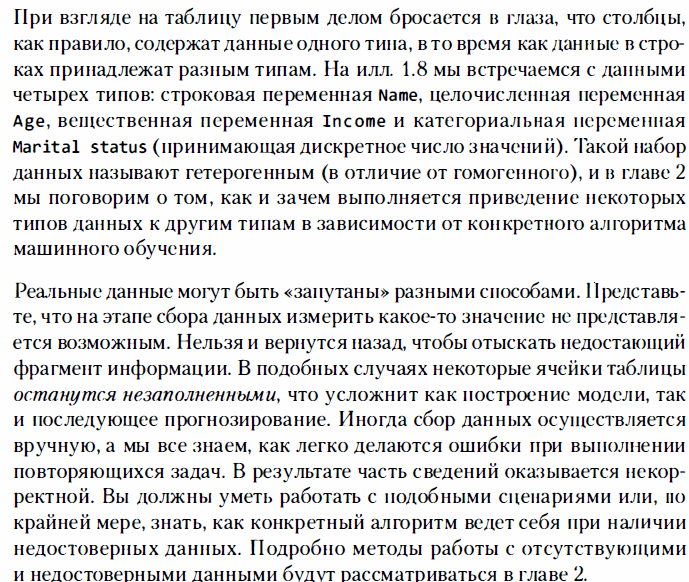
1. **Базовый процесс интеграции моделей машинного обучения в приложения или конвейеры данных.**

Его можно разбить на пять стадий: подготовка данных, построение модели , оценка, оптимизация и прогноз на новых данных.

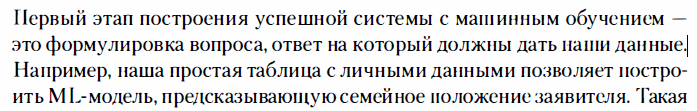
1. **Сбор и подготовка данных для систем с машинным обучением.**

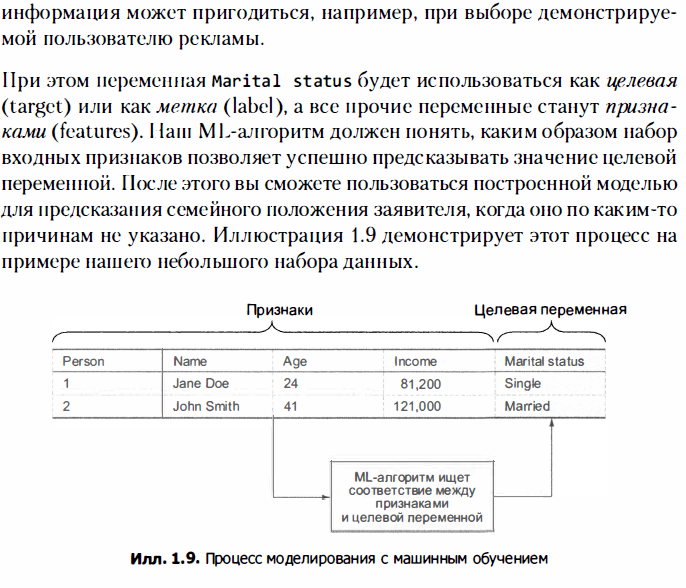


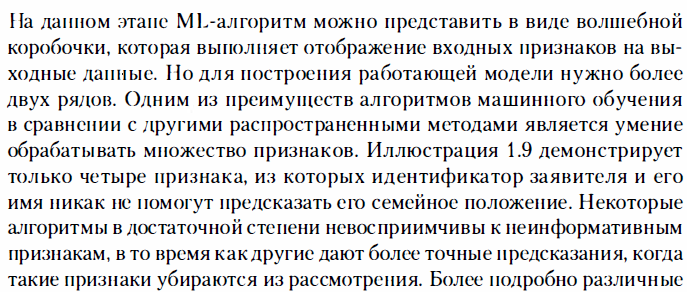




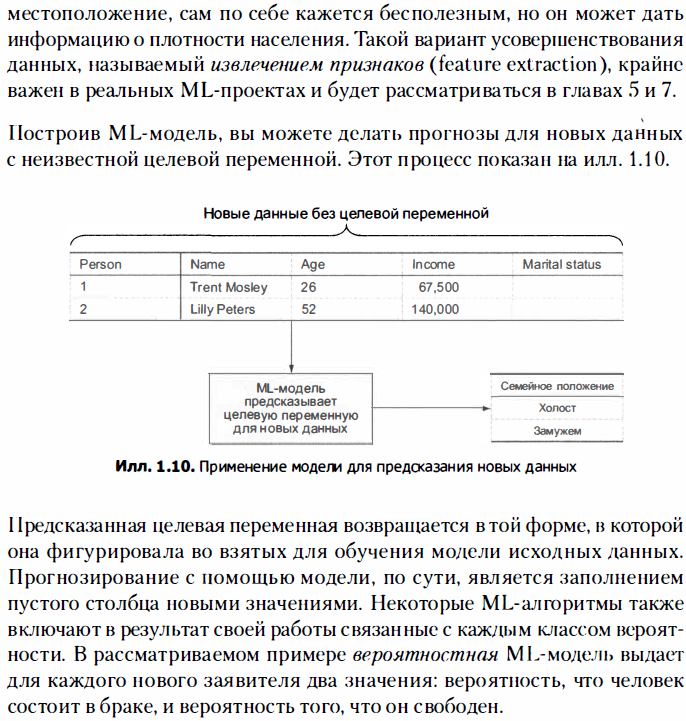
1. **Обучение модели на данных.**

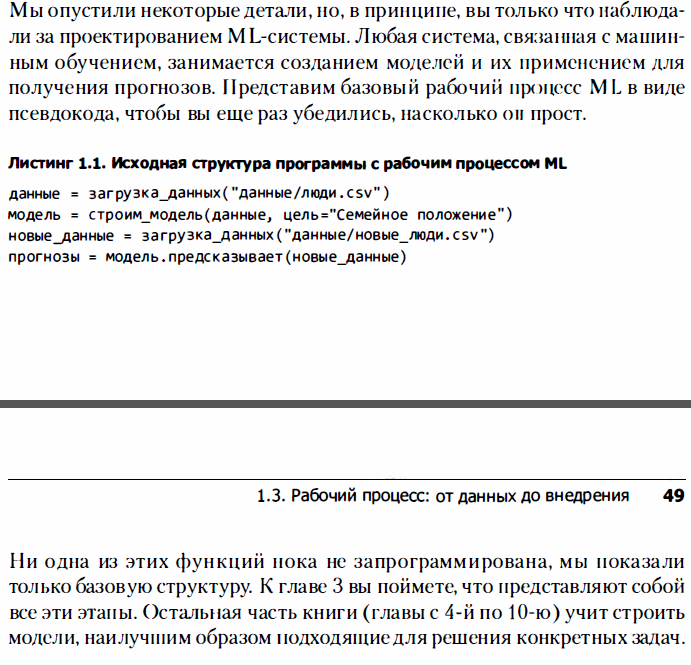












1. **Оценка производительности модели.**

c. 49

1. **Оптимизация производительности модели.**

50

1. **Предварительная обработка данных и проектирование признаков в моделях машинного обучения.**

51

1. **Непрерывное совершенствование моделей машинного обучения. Масштабирование моделей.**

54

1. **Определение набора входных признаков в задачах машинного обучения.**

62

1. **Наблюдаемое значение целевой переменной в задачах машинного обучения.**

64

1. **Определение размера обучающей выборки для построения и запуска модели машинного обучения.**

65

1. **Репрезентативность обучающей выборки при построении моделей машинного обучения.**

68

1. **Категориальные признаки в ML-наборах данных. Преобразование категориальных признаков в численные.**

70

1. **Подходы к работе с отсутствующими данными.**

73

1. **Нормализация данных для ML-алгоритмов.**

78

1. **Визуализация данных в задачах машинного обучения.**

80

1. **Поиск связи между входными данными и целевой переменной.**

93

1. **Параметрические модели с машинным обучением.**

97

1. **Непараметрические модели с машинным обучением.**

98

1. **Обучение с учителем и без учителя в задачах с машинным обучением.**

100

1. **Задача классификации. Построение классификатора и получение предсказаний.**

103

1. **Классификация сложных нелинейных данных.**

108

1. **Классификация в случае множества классов.**

111

1. **Регрессия. Построение регрессора и генерация прогнозов.**

115

1. **Регрессия для сложных нелинейных данных.**

119

1. **Оценка прогностической точности модели на новых данных.**

125

1. **Сущность проблемы переобучения ML-модели.**

125

1. **Оценка точности модели. Метод отложенных данных.**

129

1. **Оценка точности модели. Контроль по k-блокам.**

131

1. **Оценка моделей классификации.**

135

1. **Компромиссы при оценке точности моделей и ROC-кривые.**

140

1. **Многоклассовая классификация.**

144

1. **Оценка моделей регрессии.**

147

1. **Оптимизация ML-модели путем подбора параметров.**

152

1. **Сеточный поиск как способ оптимизации ML-модели.**

154

1. **Причины проектирования признаков. Основные этапы проектирования признаков.**

163, 166

1. **Прямой отбор и обратное исключение признаков.**

178

1. **Отбор признаков для исследования данных.**

180