

Билет 2

Щербакин С.Е.

71-ПГ

1 Задачи, решаемые с применением машинного обучения.

Машинное обучение можно применить в широком спектре задач. Тут, наверное, стоит сказать, где машинное обучение применить не можно.

На данный момент, это только люди не знают. Хотим ли определить кошку или собаку - пока нельзя.

Хотим ли получить точное предсказание где-то в экономической сфере. - тоже пока нельзя.

Также, нельзя не отметить большое вклад ученых в медицинскую область. Как пример, в прошлом году теньюна научились



награду за вклад в изучение  
рассеянной монеты.

Еще одним ярким примером  
монеты выступают кредитные споры.

Вместо гниющего и монотонного  
человеческого труда, машина определяет  
его - дать человеку кредит или не  
дать, основываясь на его кредитной  
истории.

Также, машинное обучение  
хорошо себя показывает в задачах,  
где нужно проанализировать большое  
количество признаков.

Разумеется, для всего этого нужно  
большое количество данных. Поэтому, на  
их подготовку уходит большая  
часть времени.

Основными типами задач  
машинного обучения являются:



- Классификация - на основе данных ~~то~~ необходимо определить класс. Примеры: кот или собака, болен человек или нет.
- Регрессия - на основе данных необходимо предсказать численное значение. Пример: температура воздуха.
- Рекомендации - альтернатива для человека, которая ему понравится. Пример: видео на YouTube.
- Заполнение пропусков - Заполнение данных на основе существующих. Пример: экономические модели.



Бизнес 2

Учебники С

71-ПГ

2. Обучение с учителем и без  
учителя в задачах с машинным  
обучением.

В общем и целом, задачи  
машинного обучения можно разделить  
на 2 типа: обучение с учителем  
и обучение без учителя.

Большая часть алгоритмов  
машинного обучения относится  
к первому типу. Это и линейная  
регрессия, и дерево решений, и  
линейный классификатор и многие другие.  
Суть данных методов заключается  
в том, что у нас уже заранее  
есть данные для обучения, в которых  
определена целевая переменная.

Однако, встречаются также наборы  
данных, в которых целевая переменная



осуществляется. Вот тут на помощь и приходят алгоритмы обучения без учителя.

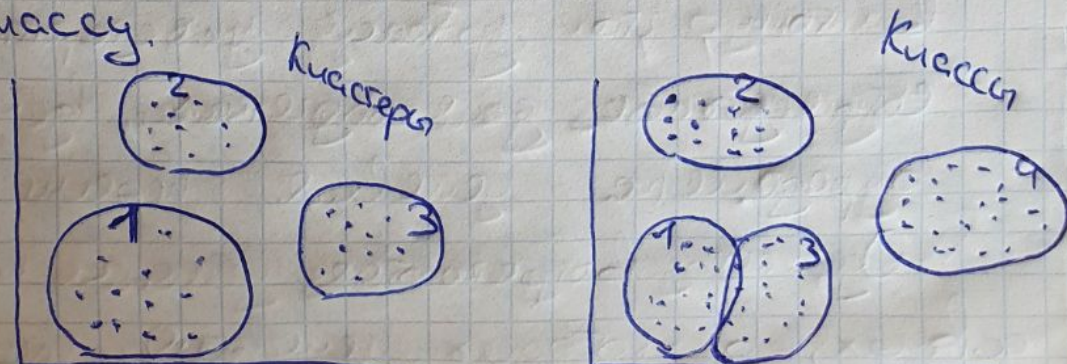
Их можно разделить на 2 основных категории:

- кластеризация данных

Здесь набор входных данных (признаки) используется для нахождения точек пересечения и последующего разбиения данных на группы (кластеры)

Примеры алгоритмов: k-средних, иерархическая кластеризация.

Суть состоит в том, что изначально данные принадлежат одному кластеру в действительности принадлежат одному классу.





- сокращение размерности

Здесь входные данные преобразуются в небольшое количество координат, при этом в них сохраняется информация о самом объекте.

Примеры алгоритмов: метод главных компонент,

Оба подхода к обучению себя не конкурируют, а наиболее популярными алгоритмами являются метод  $k$ -средних и метод главных компонент.

Проблемой тут выступает тот факт, что некоторые программы не могут работать с данными и вводить зависящую переменную, и поэтому прибегают к обучению без учителя там, где это не нужно.

Также, методы обучения без учителя используются вместе с обучением без учителя.

Это позволяет добиться более  
точной модели, за счет введения  
новых зависимостей и признаков.