



Physical Activity Clustering

Над проектом работал ученик 7 класса
Ледовских Марк

задача

В современном мире существуют множество различных устройств (фитнес-браслетов, камер, датчиков и т.д.), которые собирают данные об активности пользователя. Эти сведения могут быть использованы, например, для отслеживания состояния здоровья человека и определение находится ли он сейчас в безопасности, что он делает и так далее.

Данные, собранные с носимых устройств, которые отслеживают физическую активность пользователей, используются во множестве коммерческих приложений, включая здравоохранение, фитнес-тренировки, навигационную систему и так далее.

В данной задаче вам предлагается, используя информацию, собранную с помощью трех Инерциальных измерительных модулей (IMU) и пульсометра разработать модель машинного обучения.



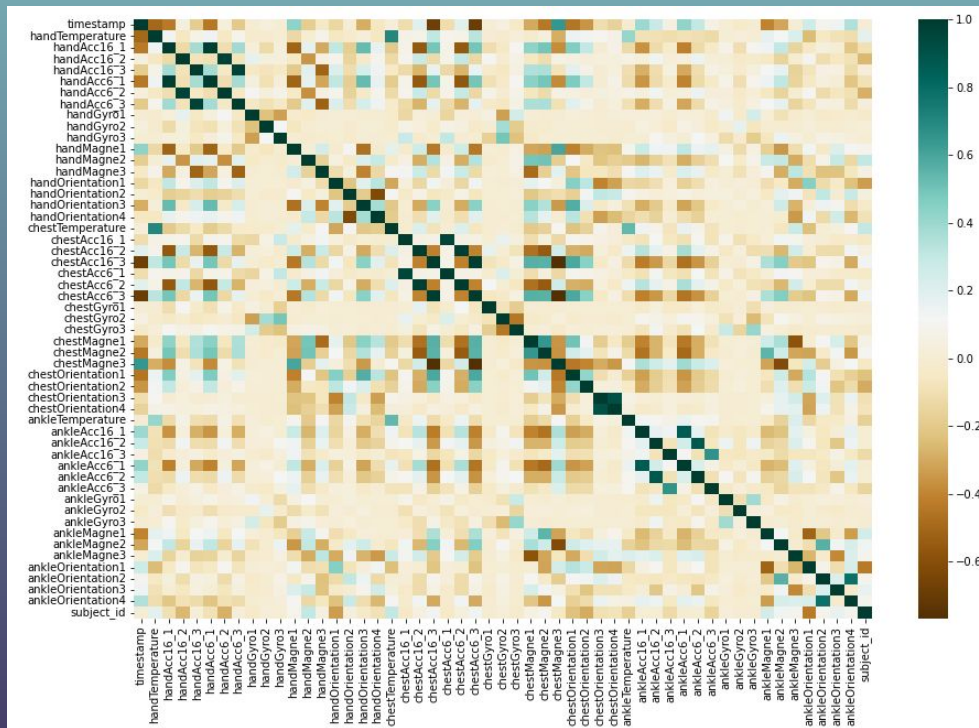
Этапы

Задачу я разделил на несколько этапов

1. Анализ данных
2. Обработка данных (удаление ненужных столбцов и заполнение пропусков)
3. Использование разных методов обучения
4. Поиск наилучшего результата

Анализ данных

В начале я проанализировал пропуски в данных и построил корреляционную таблицу.



Обработка данных

Я удалил столбцы в которых было слишком много пропусков. В тех же столбцах где пропуски были, но их было меньше заменил на средние, медианные и использовал интерполяцию. В будущем я хотел сравнить все эти способы по результатам.

Использование разных методов обучения

Используя несколько методов оказалось, что лучший метод для этой задачи - это метод локтя. После этого я попробовал поменять способ заполнения пропусков и лучшим был заполнение средним значением.

Результаты

Максимальный достигнутый результат 0.38141 / 1 балла.

Итоги

- Создана модель, которая делит на группы датчики измерения.
- лучший способ обучения - метод локтя.