

Sensing situations for smart entertainment

Andreas Jacobsen (anjac@itu.dk), Thomas Blichfeldt (tbli@itu.dk)



Problem

Recommending media for a room can be difficult without knowledge of the room.
We want to build a system, that is able to understand a specific setting within a room.

Sensors

The setting or situation is based on numerous variables like noise, temperature, light, etc.



Approach

A prototype using sensors to collect data and machine learning for classifying room settings.

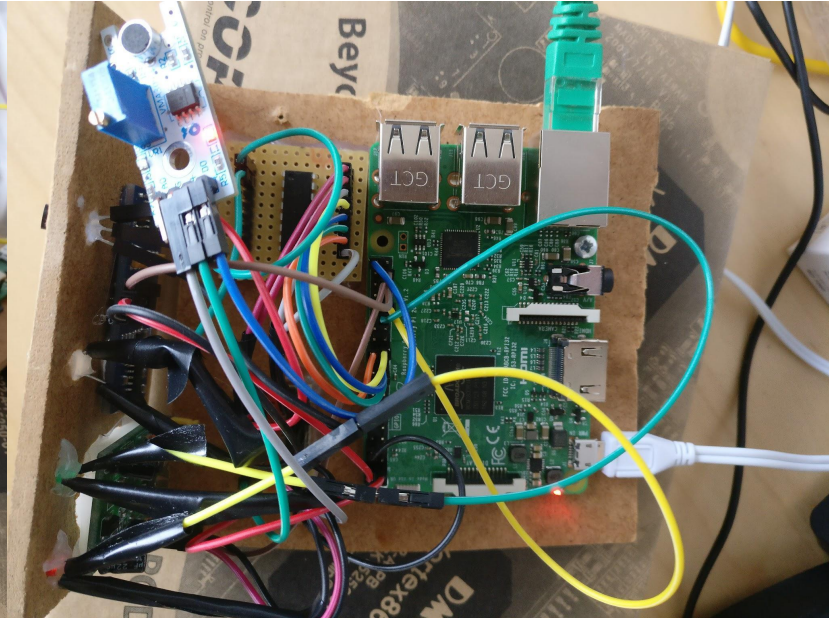
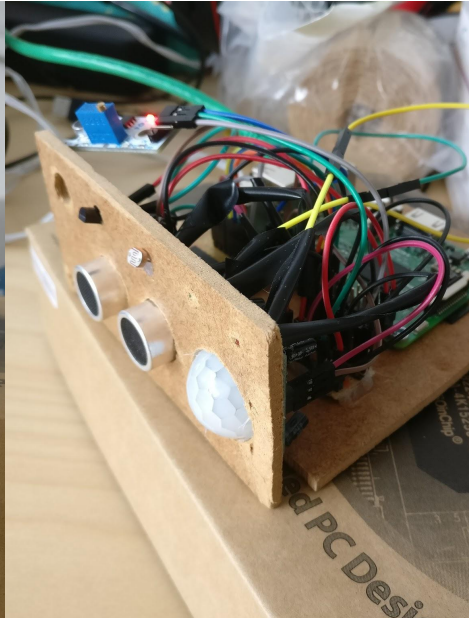
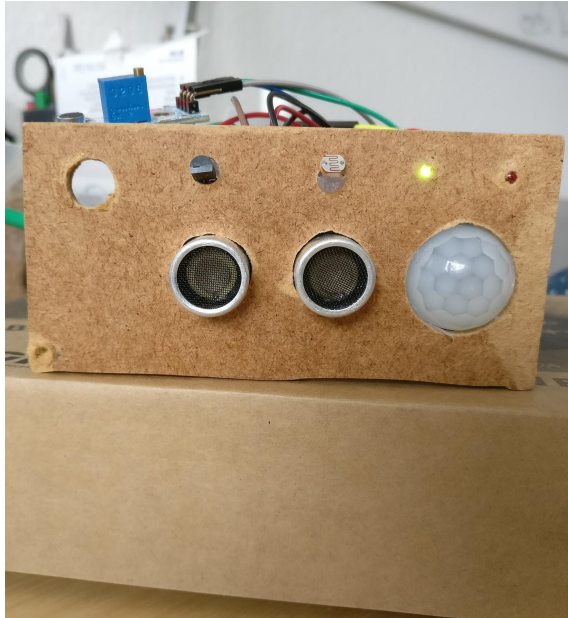
Minimal User Interaction

The system should require minimum active interactions from the user, and not require the possession of smartphones or smartwatches.

Privacy by Design

The system should not collect any sensitive data, like sound or images, that can be used for recognizing faces, locations, objects or voices.

Prototype





Experiments

Collection of Data

Three categories of data: quiet single-person event, quiet multi-person event, party multi-person event.

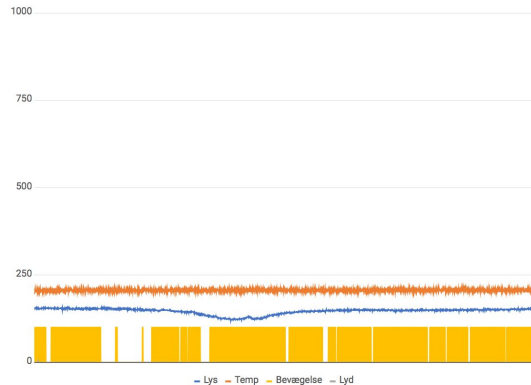
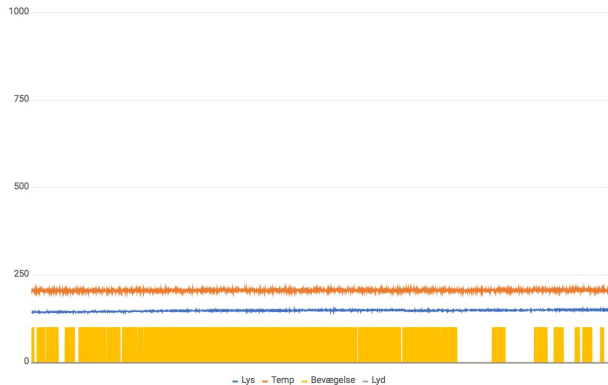
Over 50.000 data points, in the three categories.

Testing data and data relevance on the existing test set.

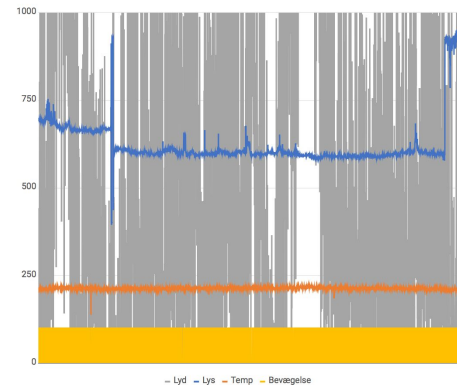
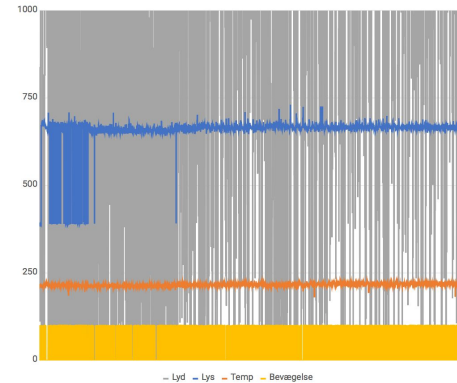
Live testing our model in the different situations.

Examples of collected data

Collected from single-person quiet events and from multi-person party event.

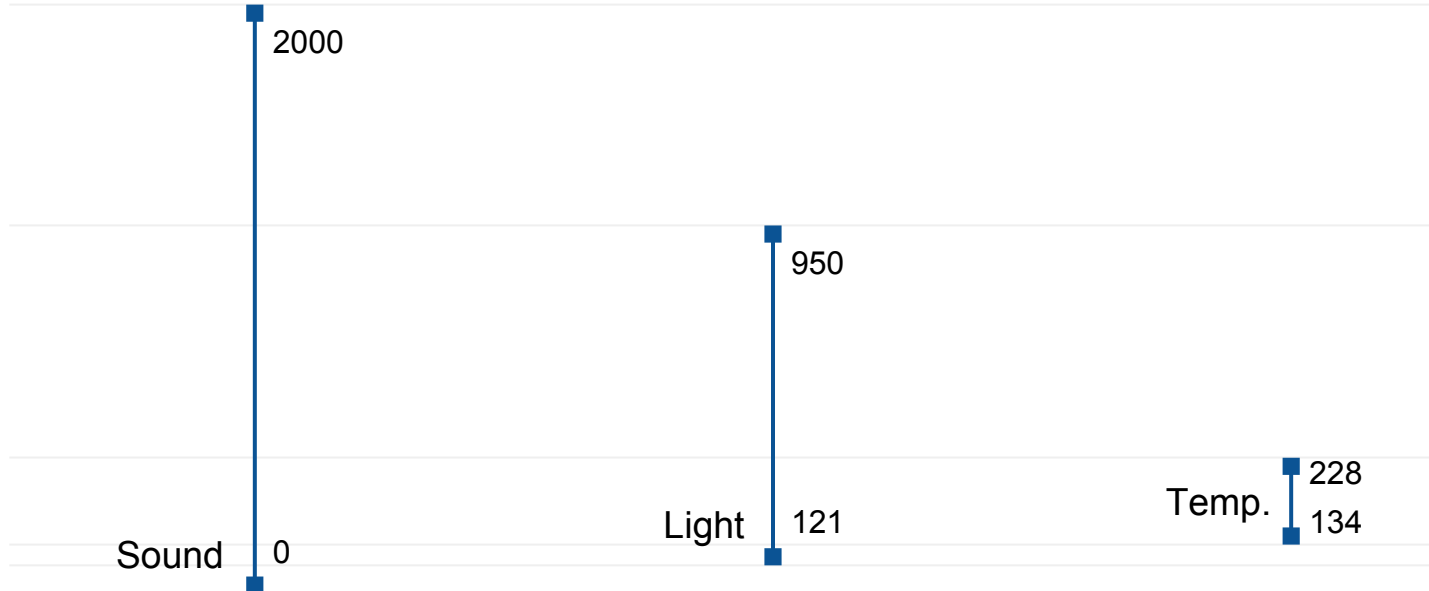


Single-person, quiet events



Multi-person, party events

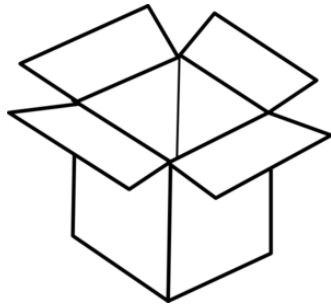
Collected data minimums and maximums



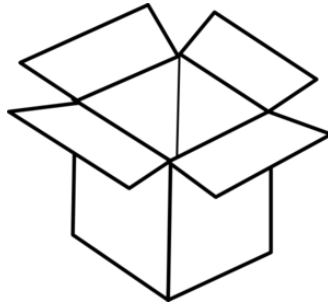


Machine Learning Classifications

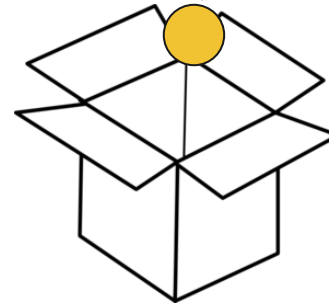
Supervised classifications, with three possibilities



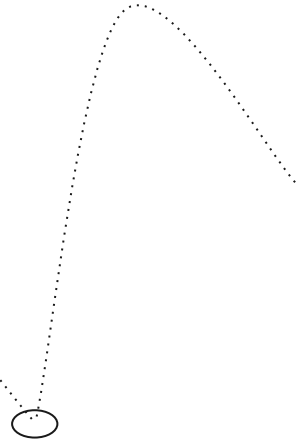
Single-person, quiet event



Multi-person, quiet event

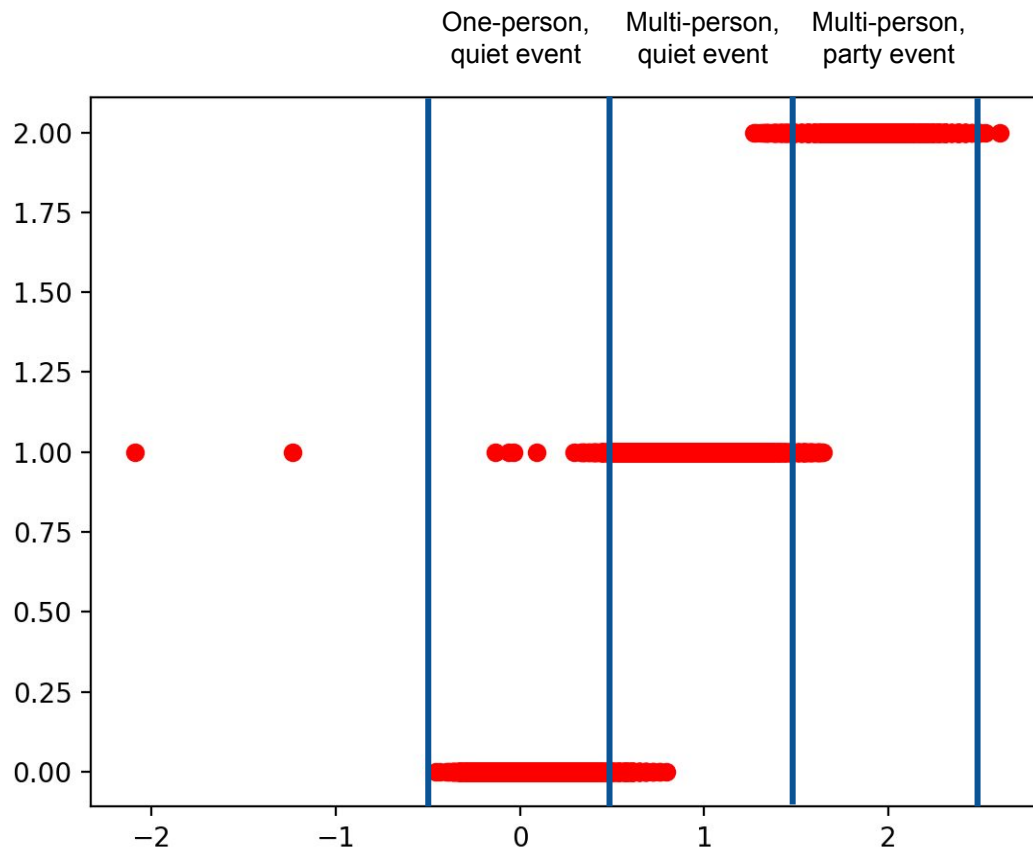


Multi-person, party event





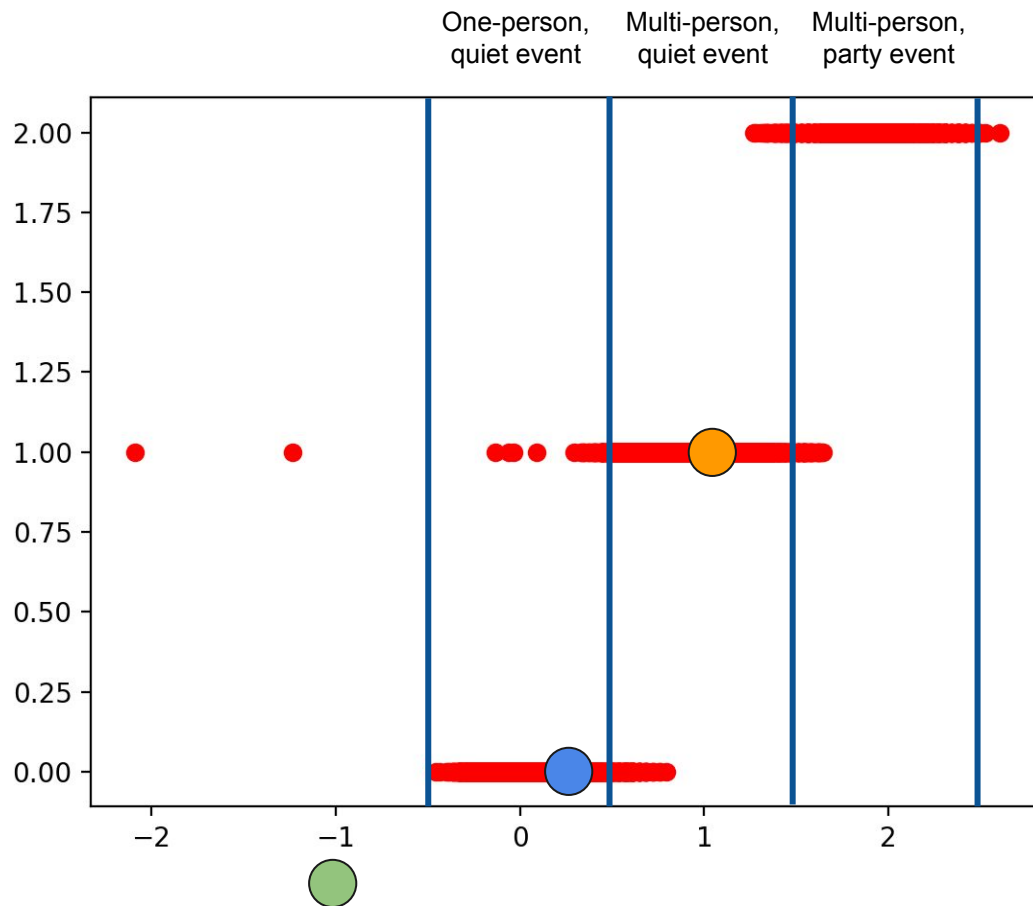
Overview of classified data points, separated into the three main categories.



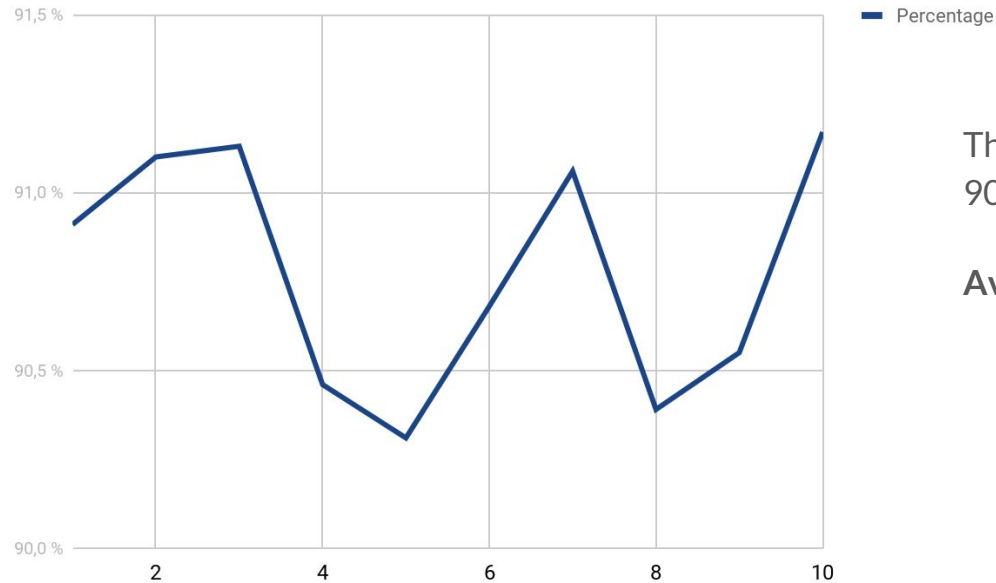
Testing

Tested Home and Social situation:

- Social averaged around 1.1
- Home averaged around 0.3
- Unknown rooms averaged around -1



10-fold cross validation



The 10 results vary from
90,3 % to 91,1 %

Average: 90,77 %



Concerns

- 10-fold cross validation gives a very high average.
- Machine learning can loop itself into bad classifications
- Few categories and small data set
- Sound should have a higher significance, but light and temperature seems more important
- Inconsistent or bad sensors
- Room over situation?

287, 211, 400, 0, 491.541624069, 17, 2
274, 210, 1394, 0, 455.559492111, 17, 2
286, 209, 1039, 0, 493.711233139, 17, 2
274, 213, 388, 1, 115.640163422, 17, 2
269, 209, 580, 1, noRead, 17, 2
265, 209, 1933, 1, noRead, 17, 2
267, 206, 1325, 0, noRead, 17, 2
315, 208, 1153, 0, noRead, 17, 2
351, 212, 748, 0, 140.624046326, 17, 2
364, 207, 925, 0, 144.137144089, 17, 2
303, 206, 866, 0, 139.998197556, 17, 2
275, 212, 1595, 1, 83.9304924011, 17, 2
266, 208, 603, 1, 123.233795166, 17, 2
302, 207, 744, 0, 88.0944728851, 17, 2
312, 210, 749, 0, 125.620365143, 17, 2
309, 208, 701, 0, 109.406709671, 17, 2
267, 210, 1082, 0, 491.466522217, 17, 2
263, 209, 1303, 0, noRead, 17, 2
264, 209, 1226, 0, noRead, 17, 2
263, 209, 162, 0, noRead, 17, 2
263, 212, 641, 1, noRead, 17, 2
264, 210, 808, 1, noRead, 17, 2
263, 210, 107, 0, noRead, 17, 2
263, 209, 326, 0, noRead, 17, 2
260, 211, 0, 0, noRead, 17, 2
263, 209, 228, 0, noRead, 17, 2
263, 210, 32, 0, noRead, 17, 2
262, 207, 481, 0, noRead, 17, 2
264, 211, 91, 1, noRead, 17, 2
263, 210, 821, 1, noRead, 17, 2
264, 208, 307, 0, noRead, 17, 2
264, 208, 80, 0, noRead, 17, 2
264, 206, 0, 0, noRead, 17, 2
265, 211, 23, 0, noRead, 17, 2
263, 210, 0, 0, noRead, 17, 2
266, 211, 38, 1, 511.034727097, 17, 2
264, 210, 0, 1, 521.190166473, 17, 2
264, 208, 31, 1, noRead, 17, 2
265, 206, 171, 0, noRead, 17, 2
265, 207, 143, 0, noRead, 17, 2
266, 205, 227, 0, noRead, 17, 2
279, 210, 0, 0, noRead, 17, 2

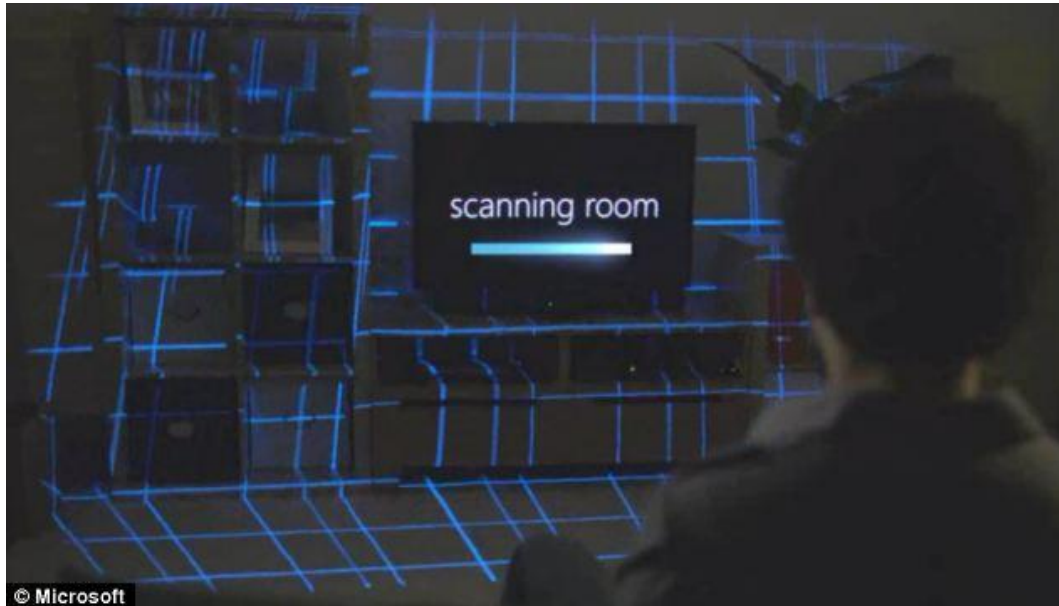
Data

noRead - The distance were not captured. This is replaced by the average, but can still give false readings.

Noise varied from 0 to 2000. Even when the party and the music were consistently high.

Movement has a lot of LOW readings even though there were a lot of movement.

Future Work



Learn new classification by itself, based on the user interactions from the recommendation service.

E.g. learn to play christmas songs, when it is a quiet multi-person event in December.

Have more devices for a broad understanding of the room