kaggle   Search kaggle

Competitions   **Datasets**   **Kernels**   **Discussion**   **Jobs**   Sign

🏆 Featured Dataset

6

The ExtraSensory Dataset

Behavioral Context Recognition In-the-Wild

Yonatan Vaizman • last updated 5 months ago

Overview   Data   Kernels   Discussion   Activity   Download (8 MB)   **New Kernel**

Tags   psychology   medium   featured

Top Contributors		
Yonatan Vaizman	1st	
Austin	2nd	
Kristofer Engelbrek...	3rd	

Kernels		
Reading user data and basic ...	1 vote	
run 5 months ago		
Can user behavior be determi...	0 votes	
run 5 months ago		
testing	0 votes	
run 5 months ago		

Discussion

There are no conversations yet.

Start one

Description

Context

Behavioral Context refers to a wide range of attributes describing what is going on with you: where you are (home, school, work, at the beach, at a restaurant), what you are doing (sleeping, eating, in a meeting, computer work, exercising, shower), who you are with (family, friends, co-workers), your body posture state (sitting, standing, walking, running), and so on. The ability to automatically (effortlessly, frequently, objectively) recognize behavioral context can serve many domains. Medical applications can monitor physical activity or eating habits; aging-at-home programs can log older adults' physical, social, and mental behavior; personal assistant systems can better server the user if they are aware of the context. In-the-wild (in real life), natural behavior is complex, composed of different aspects, and has high variability. You can run outside at the beach, with friends with your phone in the pocket; you can also run indoors, at the gym, on a treadmill, with your phone motionless next to you. This high variability makes context-recognition a hard task to perform **in-the-wild**.

Content

The ExtraSensory Dataset was collected from 60 participants where each person participated approximately 7 days. We installed our data-collection mobile app on their *personal phone* and it was used to collect both sensor-measurements and context-labels. The sensor-measurements were recorded automatically for a window of 20-seconds every minute. This included accelerometer, gyroscope, magnetometer, audio, location, and phone-state from the person's phone, as well as accelerometer and compass from an additional smartwatch that we provided. In addition, the app's interface had many mechanisms for self-reporting the relevant context-labels, including reporting past context, near future, responding to notifications, and more. The flexible interface allowed to collect many labels with minimal effort and interaction-time, to avoid interfering with the natural behavior. The data was collected in-the-wild: participants used their phone in any way that was convenient to them, they engaged in their regular behavior and reported an combinations of labels that fit their context.

For every participant (or "user"), the dataset has a CSV file with pre-computed features that we extracted from the sensors and with labels. Each row has a separate example (representing 1 minute) and is indexed by the timestamp (seconds since the epoch). There are columns for the sensor-features, with the prefix of the column name indicating the sensor it came from (e.g. prefix "raw_acc:" indicating a feature came from the raw phone accelerometer measurements). There are columns for 51 diverse context-labels and the value for an