Milestone 1

Team Name
Urban Analysts

Team Members
Angelina Paredes
Bryton LeValley
Taryn Burns
Melanie Bell

Part 1 Project Description

1. Problem Statement

1.1. What's the input and output of the problem?

1.1.1. We're going to use sensor data that monitor different activities in a household. The data collected from the smart homes are motion, door, item, light, temperature, battery, zigbee, others, and misc. Entries include in the data are date, time, sensor name, message, and sensor type. From our data, we are going to find out the most efficient strategy to sustain a smart home. We'll also find a frequent sequence of sensors that a resident triggers in their home. Using testing and training data, we'll implement an activity recognition program.

1.2. Why the problem you want to address is important? What's the application?

1.2.1. Activity tracking will revolutize independent living for diasabled and elderly individuals. With sensors, professionals can monitor the activity of residents and ensure they're living an active lifestyle. Testbeds make sure that the sensors are accurate and efficient. Our application is an activity analyzer and sensor monitor using testbed data. The application will display information on sensors and household activity.

1.3. Why is it hard? Where is the challenge?

1.3.1. There is one challenge that can be discussed as a major point. When there are people active in the house, how will we tell the sensors that they're still active? For example, the people in the home are currently watching a movie or TV shows on a streaming site on the TV. They're technically active but not fully active. We want our sensors to be aware that even though there are people in the room, we have the sensors be idle but be aware that they may be used at any moment so they do not have to be fully turned off as if someone wasn't home.

We will end up running into some challenges throughout the project but with the more we progress through the semester, we will be able to figure out what each challenge will bring.

1.4. Specify the goal you want to achieve (a new problem, improvement for algorithms, and/or a thorough experimental evaluation of existing algorithms).

1.4.1. The collective goal our team is to design an unsupervised learning algorithm to classify unlabeled data. This will allow us to analyze as much data as possible in order to get the best model of daily living (Cook, et.al 2013).

As a whole, our team has a variety of individual goals that we have set for this project. A goal that Melanie has, is to gain a greater understanding in building a database from scratch. In specific, building a database that is intuitive and user-friendly for those who may not understand the data as indepthly as the creators, or managers of the database.

Angelina's personal goal is to learn more about large datasets and how to utilize them to one day better the daily lives of others. Technology is constantly changing today, which means that data overall is constantly evolving and growing. Over the last ten years, our homes have grown into smart homes where certain technologies can control practically anything, even if it is just a voice command.

1.5. Your team: how many students? If more than two students, you should give a justification on each team member's responsibility: who solves what sub-problems in question d.

1.5.1. Our team consists of four different students. At this point in time, we have not delegated any specific work to any individual. We are still in the process of planning and figuring out how we are going to approach the problem. However, one person will be in charge of the Github repository even though all of us will be on it. Two others will be in charge of the database and finally one person will be in charge of turning in our Milestone documents.

2.   Datasets and Tools

2.1.   Give the link and description of the dataset

2.2.   Link to Dataset:

https://drive.google.com/file/d/1QW_EPdGPDgikKFy9xrGSB_sjJ6YF9hUH/view?usp=sharing

Our dataset is sensors collected from a two resident apartment in 2009-2010 academic year. The residents are labeled as R1 and R2. They were motion, door, kitchen items, burner, hot water, cold water, temperature, and electricity sensor data that was collected.

2.3.   Briefly state what tools you will be using to parse the data, data models you want to select

2.3.1.   We will code our analytic and parse our data in Python 3. Github will be used to share our code. We will use Amazon Relational Database Service (Amazon RDS) to create, manage, and query a postgreSQL database.

2.4.   A progress report: what have you done for preparing the data sets? What's your plan for Milestone 2?

2.4.1.   We have gathered different datasets that have been researched, and we are using CASAS datasets that are online. Since Milestone 2 is unknown, we are unsure of what will come next but we are continuously researching and planning out the datasets we will be using. Additionally, after discussing various platforms, we have come to the decision that we will be using Python and PostgreSQL.

2.4.2.   We will create an accurate data model for our data in Milestone 2. Another objective is to have our data uploaded in a database and be able to run queries on it for Milestone 2.

3.   Timetable for Milestones

Dates	Milestones
9/4 - 9/15	Milestone 1
TBD	Milestone 2
TBD	Milestone 3

TBD	Milestone 4
TBD	Milestone 5
12/16	Final Project Report Due

TBD is in the slots above because we do not know when the Milestones are due at this point in time. However as a team, we will be continuously working on the project throughout the semester in order to meet the final due date.

References

Cook, Diane J., Crandall, Aaron S.,Thomas, Brian L. ,Krishnan, Narayanan C. ."CASAS: A Smart Home in a Box".2013.https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3886862/