HomeSGN: A Smarter Home with Novel Rule Mining Enabled by a Scorer-Generator GAN

Zehua Yuan, Junhao Pan, Xiaofan Zhang, Deming Chen
University of Illinois at Urbana-Champaign, IL, USA

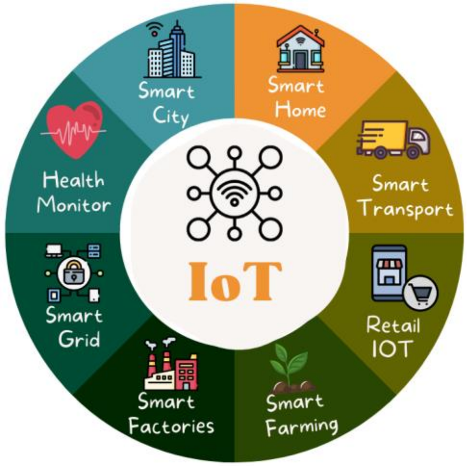
Overview

- **Background**
- **Contribution**
- **Motivation**
- **Methodology**
- **Experiment and Evaluation**
- **Conclusion**

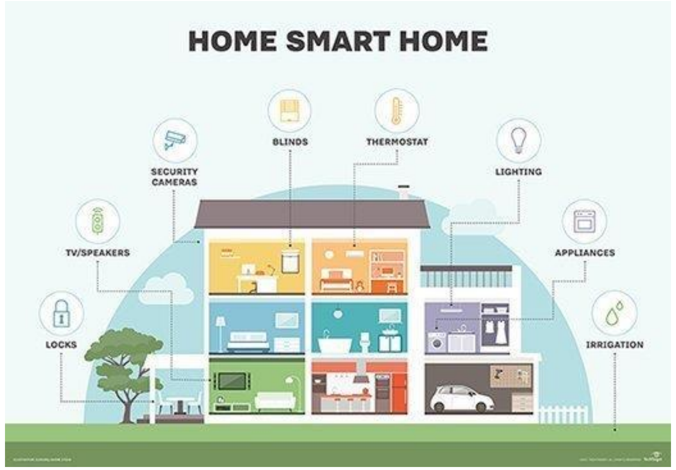
Background – IoT and Smart Home

With the advancement of edge computing technology, the Internet of Things (IoT) has become increasingly valuable in a multitude of domains.

High-performance hardware and artificial intelligence (AI) technologies further help to build a robust foundation for smart home system.



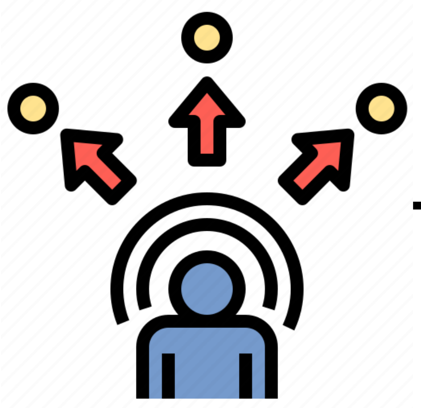
Integrate physical objects into the digital world, **enhancing efficiency** across various industries.



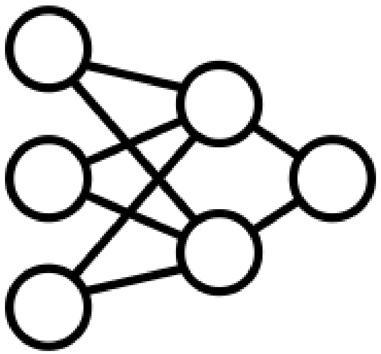
Enhance living comfort, convenience, security, and energy efficiency

Background – Deep Learning and Neural Network

Deep Learning and Neural Network: A subset of machine algorithms that mimic human brain operation, recognizing patterns and interpreting sensory data.



Input:
Data from sensors in smart homes

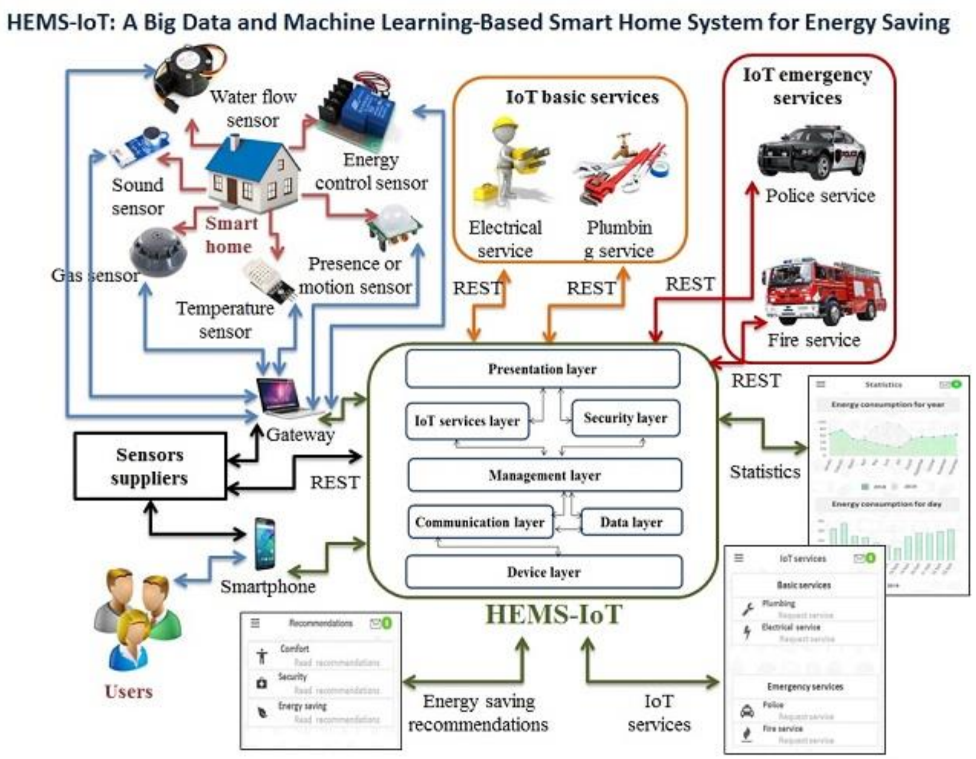


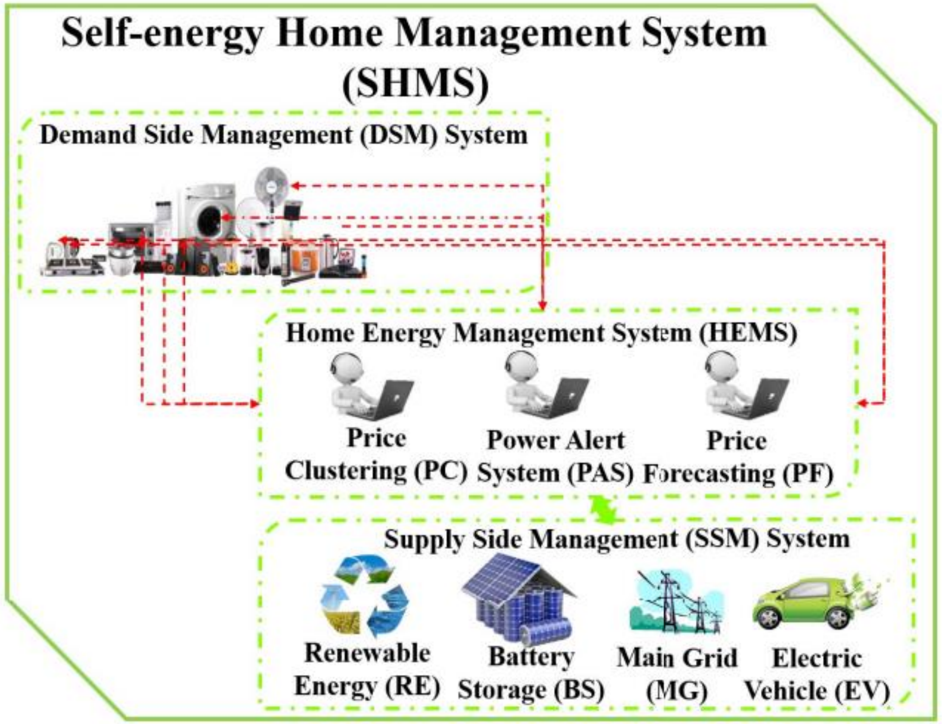
DNN

Output:
Predictions or decisions, like human activities

Broadly used in IoT and smart homes for: Energy Efficiency, Security and Surveillance, Voice and Image Recognition, Remote Control, etc.

Background – Related Work





Method: Apply J48 Machine Learning algorithm on home big data
Limitation: Focus exclusively on energy management

[I. Machorro-Cano, et al., Energies, 2020]

Method: Apply K-means and LSTM for price forecasting, price clustering, and power alert system
Limitation: Focus exclusively on energy management

[W. Li, et al., IEEE Internet of Things Journal, 2018]

Background – Challenges

Ultimate goal:

Translate users' living patterns and environmental factors into actionable rules that could improve residents' quality of life.

Challenges:

- Complexities of household environments for data analysis
- Lack of high-quality event sequence datasets for smart home rule-mining
- Balance user-specific enhancements with household optimization

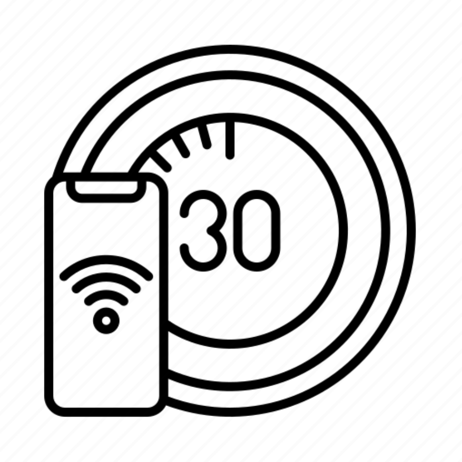
Contribution:

- HomeSGN - an innovative **generative smart home** system
  - State-of-the-art rule-mining and proposing system with scorer and generator
- Data augmentation and rectification module for data pre-processing
- Integration of smart home framework and **real-life experiment**
  - Encourage **practical optimization** for users and enhance user-home interactions

Motivation – From Device to User

- Current designs focuses on the **device level**, fails to enhance residents' living quality.
- Require user input to perform **certain functions**.
- Cannot adapt to **individual user preferences** and routines.

- Highlight the need to enhance the intelligence of these systems to **improve users' living quality.**
- Transition from a device-centric to a **user-centric** approach in smart home design.



Smart Thermostats

Learn a household's routines and **adjust heating and cooling** for optimal comfort and energy efficiency

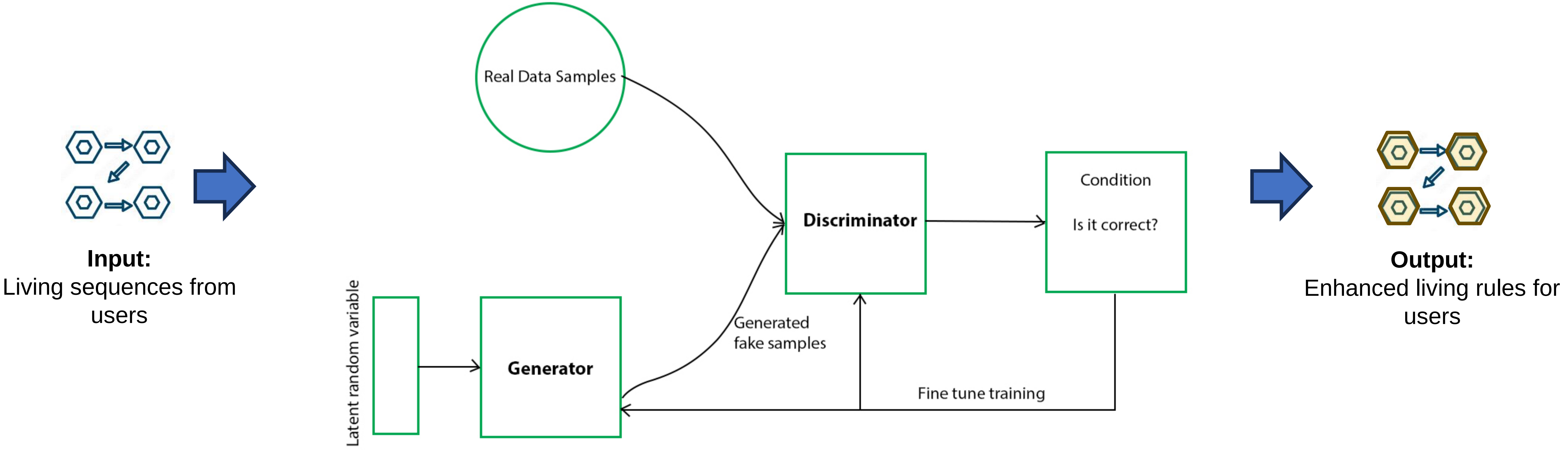
Intelligent Lighting Systems

Allow users to control lighting **remotely**, change light colors, and **set schedules** for lights to turn on or off.

Motivation – Generative Neural Network

Generative Neural Network: Neural network that generates new data samples by learning the given dataset.

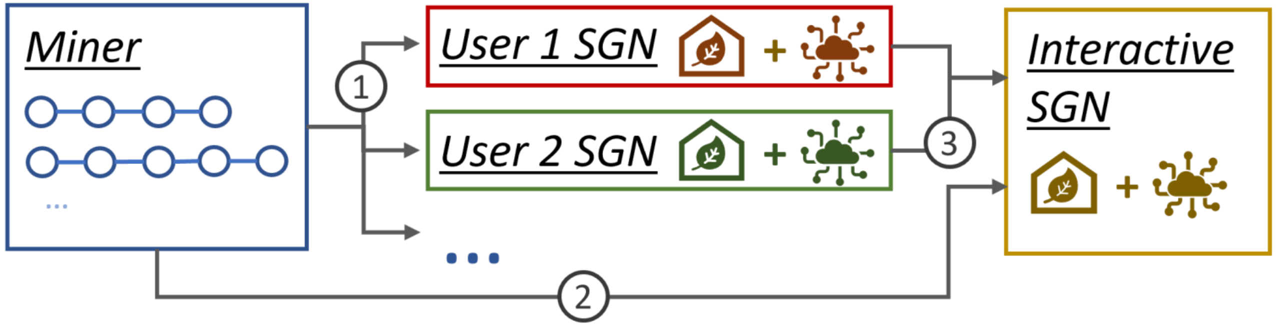
GAN (Generative Adversarial Network): A framework contains a generator and a discriminator, that compete against each other.



HomeSGN (Home Scorer-Generator Network):

Unique scorer-generator network crafted for smart home systems, targeting to propose useful new living rules based on historical living sequences.

- Miner Block: Statistic rule mining
- Scorer Block: Living rule evaluation
- Generator Block: Enhanced new rule generation



Overview of Multi-user HomeSGN

Mining Block

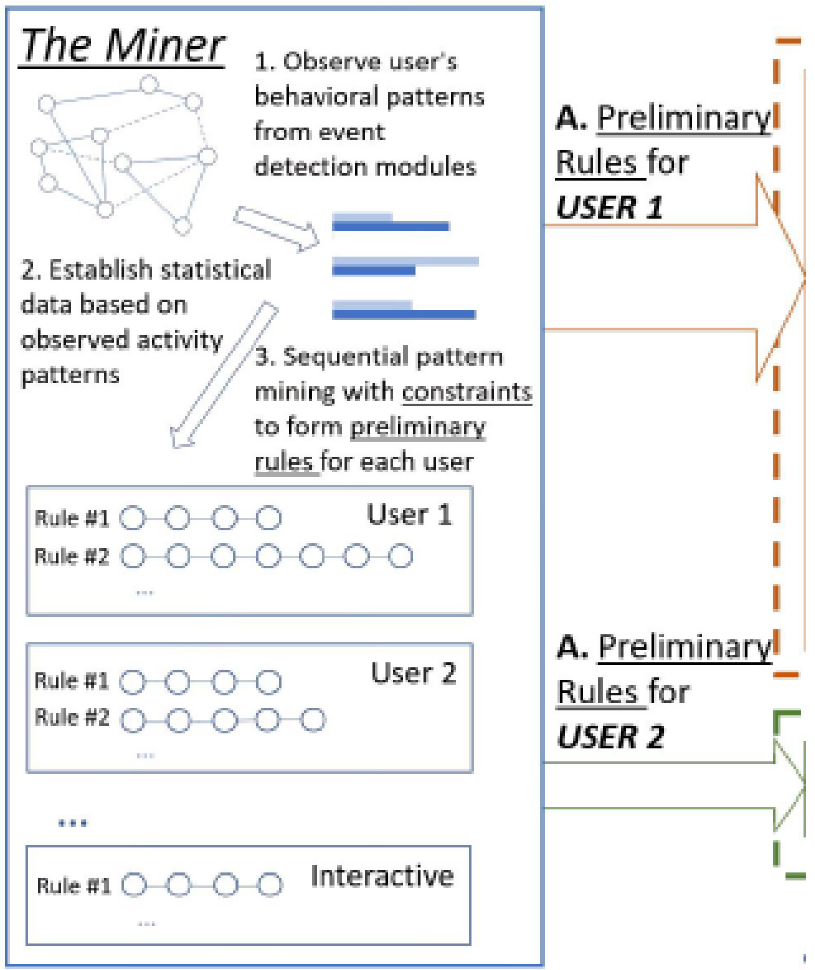
From house-level living sequence to find living patterns of individual users and interactive rules among users.

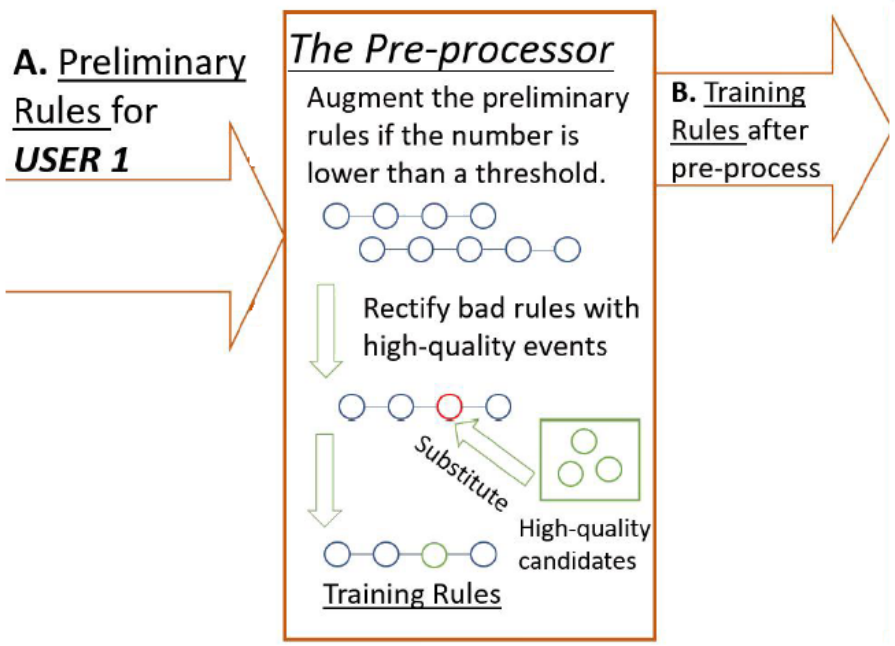
Modified MDD (Multi-valued decision diagram) algorithm:

- Input: House-level living sequences with user_ids, location_info, and timestamps.
- Output: Preliminary rules for each user, interactive rules among involved users.

Advantages:

- Smaller search space
- Enable complex and non-numerical constraints





The Pre-processor

Challenge:

- Insufficient living rules for model training.
- Training data with bad living habit.

Solution and feature of the pre-processor:

- Provide an additional good baseline for users who do not have enough training data.
- Rectify bad rules with high-quality activities in the same category to provide good living rule examples.

Methodology – Scorer Block (Encoding)

Encoding: One-hot encoding

Assuming K different activities, $Act = \{Act_i\}, 1 \leq i \leq K$, $Seq = \{Seq_j\}, Seq_j \in Act, 1 \leq j \leq L$

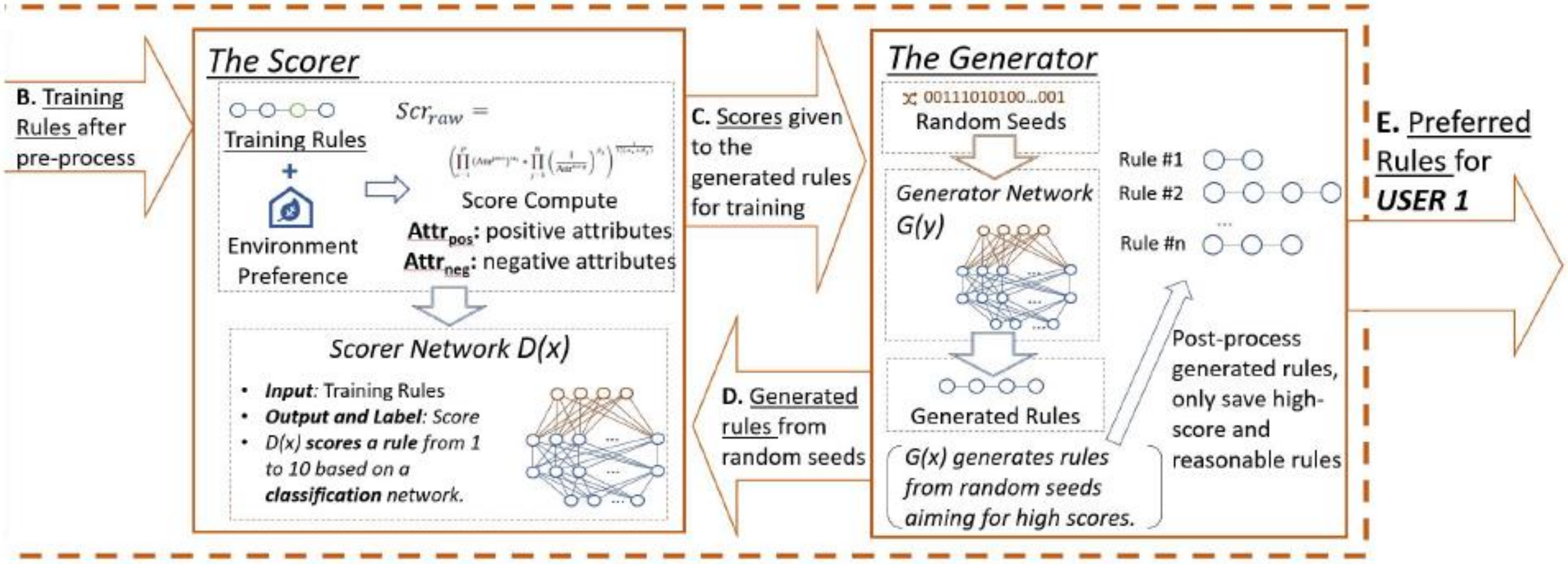
Given a certain *Seq*, we can encode *Seq* into a *L-by-K* matrix *Enc* where $Enc_{j,i} = \begin{cases} 1 & \text{if } Seq_j \text{ is } Act_i \\ 0 & \text{otherwise}x \end{cases}$

To illustrate, Assuming $Act = \{Eat, Pray, Love\}$, the sequence $Seq = \{Eat, Eat, Love, Pray\}$

is encoded as $Enc = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

Simple encoding with the values of the activities' indices, such as $Enc = [1, 1, 3, 2]$ becomes much more difficult for the neural network to learn the features **due to lack of informative data**.

One-hot encoding can also support complex activity sequences.



The Scorer-Generator:

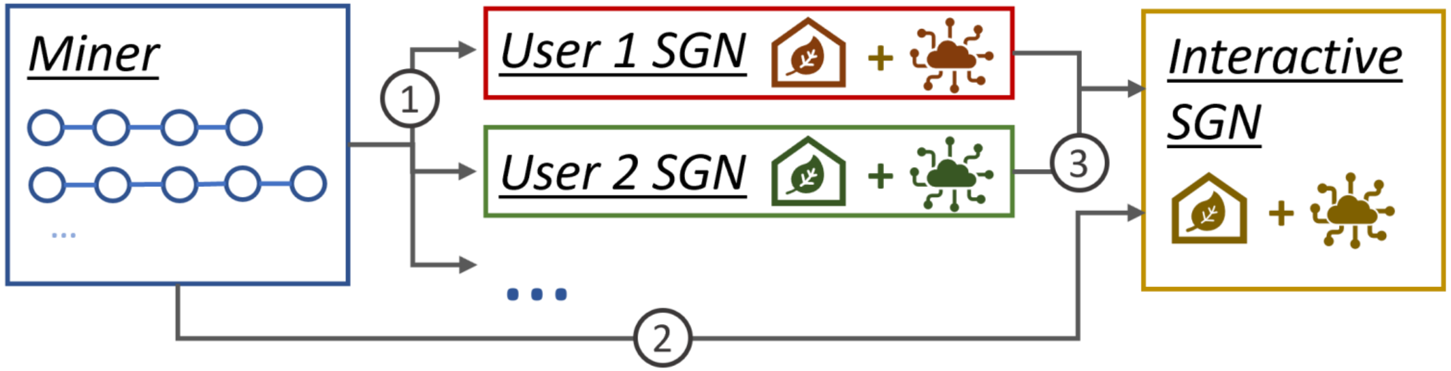
- Input: Training rules after pre-process, environment and users’ preference
- Output: Preferred Rules for each user
- Score function: $$Scr_{\text{raw}} = \left(\prod_{i=1}^{P} (\text{Attr}^{pos})^{\alpha_i} * \prod_{j=1}^{N} \left(\frac{1}{\text{Attr}^{neg}} \right)^{\beta_j} \right)^{\frac{1}{\sum(\alpha_i + \beta_j)}} ,$$
- **Attr**$^{\textbf{pos}}$ and **Attr**$^{\textbf{neg}}$ are positive and negative attributes, **alpha** and **beta** are coefficients to determine the weights of each attribute

Insights of GAN network

- High-quality generation capability
  - Remarkable ability to generate high-quality, realistic data samples
- Implicit modeling attributes
  - Implicit modeling of data distribution suits complex, multimodal data distributions

- Limitation of Variational Autoencoders (VAEs) and Diffusion Probabilistic Models
  - Cannot handle living rule data well
  - VAEs lack diversity or creativity due to their tendency to average over the training data in the latent space

Methodology – Multi-User HomeSGN



The Multi-User HomeSGN:

- Input: Sequence data for the house
- Output: Preferred rules for each user, and global optimized rules for the house
- Score function: $$Scr_{\text{raw}} = \left(\prod_{i=1}^{P} \left(\sum_{k=1}^{n} \theta_k * \text{Attr}^{pos} \right)^{\alpha_i} * \prod_{j=1}^{N} \left(\frac{1}{\sum_{k=1}^{n} \theta_k * \text{Attr}^{neg}} \right)^{\beta_j} \right)^{\frac{1}{\sum(\alpha_i+\beta_j)}}$$
- **Theta** is the coefficient to determine the privilege

Experiment and Evaluation – Experiment Setup

Experimental Setup:

- Edge node build with Raspberry Pi 4B
- Cloud node on a desktop with an Intel Xeon W-2225 CPU and an NVIDIA Quadro P2200 GPU

Dataset:

YouHome dataset[1] with profiles of two users

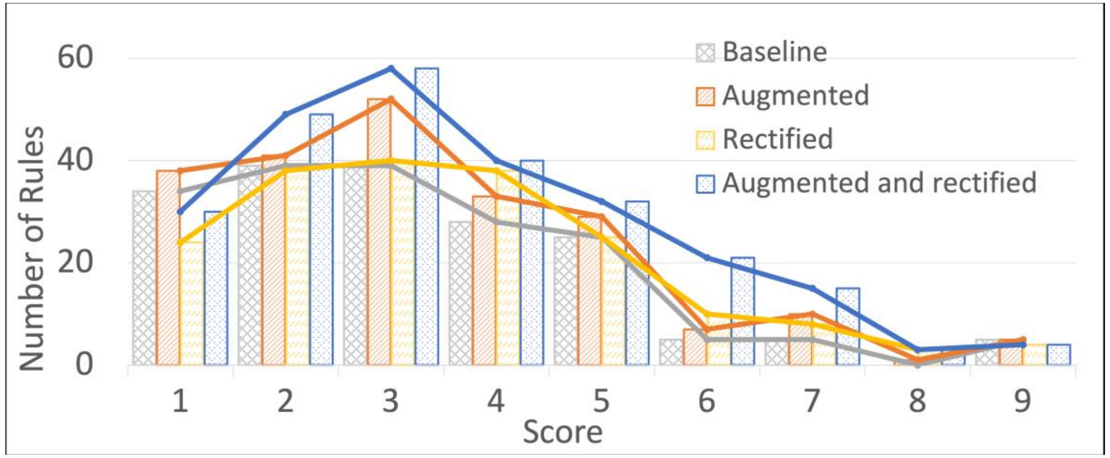
- Originally designed for activities of daily living (ADL).
- The dataset includes 45 distinct activities and 6 human postures with 20 participants in two different residential contexts with 29 camera viewpoints.
- Living sequences are derived from the dataset using sequential video data.

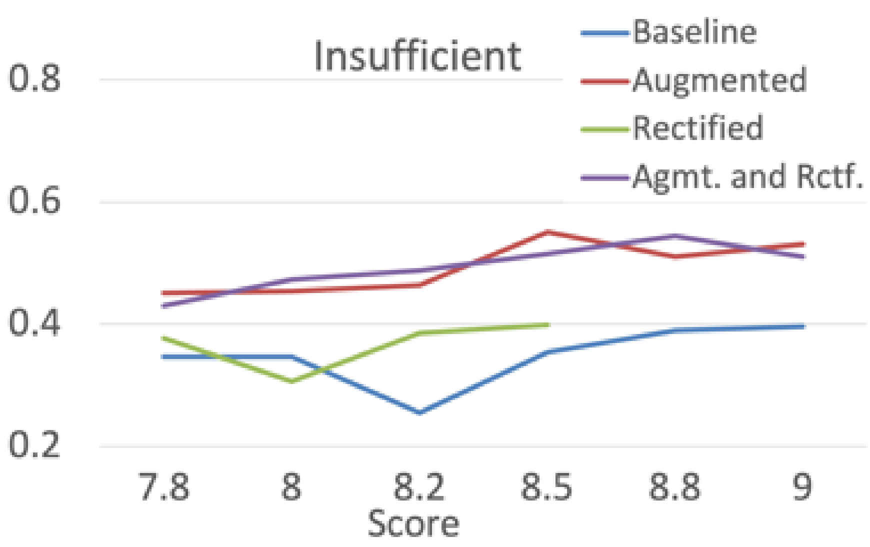
[1] The YouHome dataset is available at https://github.com/UIUC-ChenLab/YouHome-Dataset/
The repo is currently private due to paper review and will be public soon.

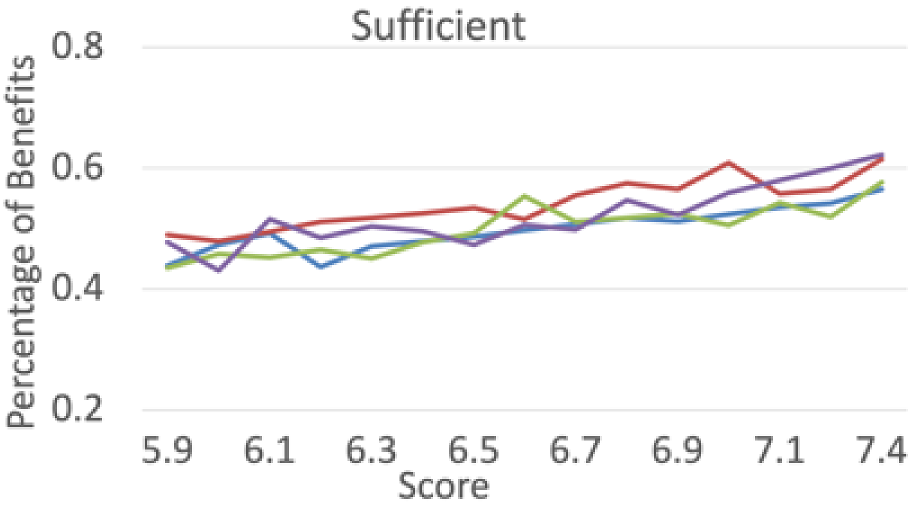
Methodology – Rule augmentation:

Rule augmentation preliminary results:

- Our method improves the number of rules in the score range of 2 to 8 compared to the baseline.
- Combining both methods significantly enhance the quality of the training data.

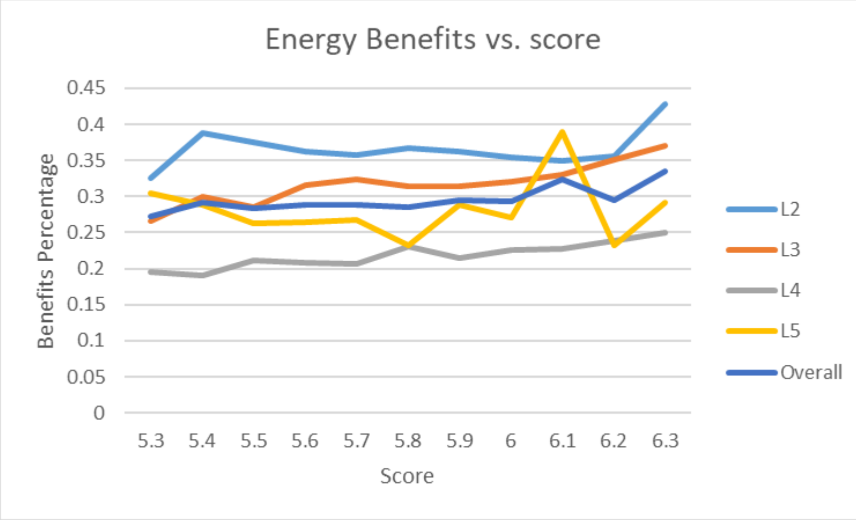


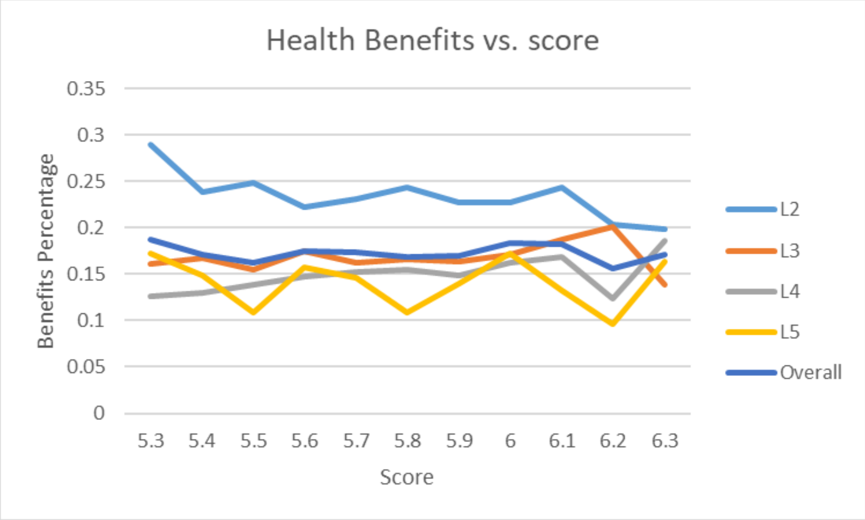


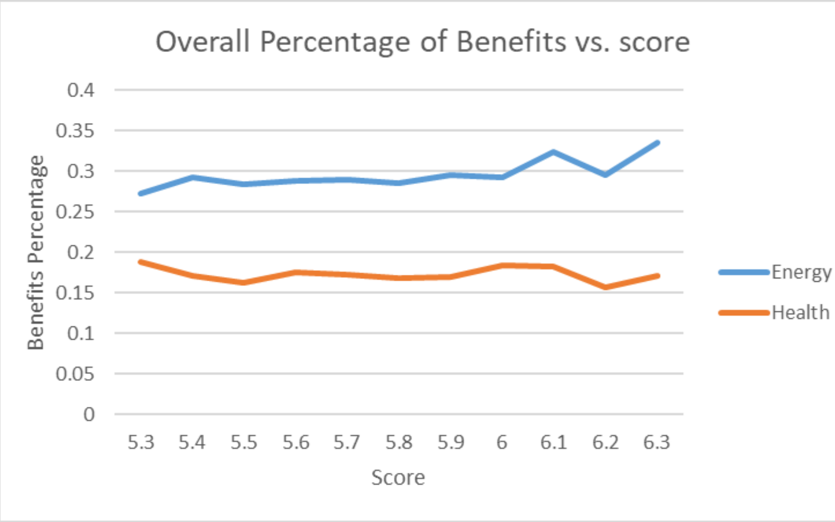


Single User Multiple Optimization Targets

- 50% Energy Factor and 50% Health Factor
- Great improvement for both attributes - 30% for energy consumption and 15% for health benefit



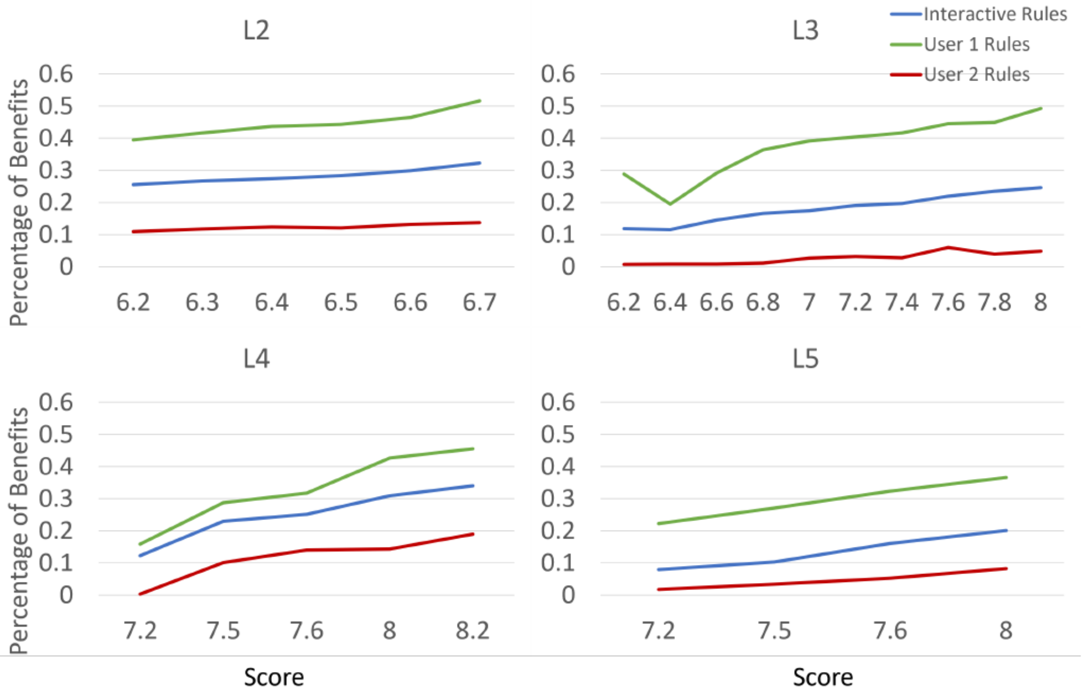




Methodology – Multi-User Multiple Optimization Targets

Multi-User Multiple Optimization Targets

- Equally optimize for Energy, Health, and Comfort
- Benefits ranging from 10% to 50% per user, 30% for interactive rules
- Interactive rule optimization could hint a user bias, require sacrifice from one user
- Introducing a privileged mode, would enable HomeSGN to more efficiently optimize for that user.



Conclusion

Novel Smart Homes:

- HomeSGN introduces novel living rule proposer for users in smart home.

Innovative Rule Mining:

- Utilizes Scorer-Generator GAN for personalized, high-value rule suggestions.

Data Augmentation and Rectification:

- Ensures high-quality results even with limited or poor-quality data.

Interactive and Adaptive:

- Offers optimized user preferences and inter-user interactive rules.

Thank You!