Synthetic Behavior Sequence Generation using Generative Adversarial Networks

FATEME AKBARI, DeGroote School of Business, McMaster University, Canada
KAMRAN SARTIPI, Department of Computer Science, East Carolina University, U.S.A
NORM ARCHER, DeGroote School of Business, McMaster University, Canada

Due to the increase in life expectancy in advanced societies leading to an increase in population age, data-driven systems are receiving more attention to support the older people by monitoring their health. Intelligent sensor networks provide the ability to monitor their activities without interfering with routine life. Data collected from smart homes can be used in a variety of data-driven analyses, including behavior prediction. Due to privacy concerns and the cost and time required to collect data, synthetic data generation methods have been considered seriously by the research community. In this paper, we introduce a new Generative Adversarial Network (GAN) algorithm, namely BehavGAN, that applies GAN to the problem of behavior sequence generation. This is achieved by learning the features of a target dataset and utilizing a new application for GANs in the simulation of older people’s behaviors. We also propose an effective reward function for GAN backpropagation by incorporating n-gram based similarity measures in the reinforcement mechanism. We evaluate our proposed algorithm by generating a dataset of human behavior sequences. Our results show that BehavGAN is more effective in generating behavior sequences compared to MLE, LeakGAN, and the original SeqGAN algorithms in terms of both similarity and diversity of generated data. Our proposed algorithm outperforms current state-of-the-art methods when it comes to generating behavior sequences consisting of limited-space sequence tokens.

ACM Reference Format:

1 INTRODUCTION

Health monitoring of older people with the aim of providing on-time care and health condition prediction has received a considerable amount of study. The availability of datasets on older people’s daily behavior can benefit a large body of studies including: applications of machine learning methods on predicting and detecting anomalous behaviors [5, 23, 27, 35, 37]; and development of reminder and recommender systems in healthcare support and the supervision of long-term behavior [7, 19, 47]. Furthermore, the efficiency and effectiveness of deep learning methods depend on the quality and quantity of training data. Due to the following reasons, existing datasets of real data do not meet the requirements of research in this area: i) scale of data: training of machine learning models tends to require large amounts of data; ii) privacy of data: health monitoring poses privacy concerns to the people whose activities are being monitored.