Improving AR/VR experiences with deep learning

Sharare Zehtabian, Siavash Khodadadeh, Ladislau Bölöni, Damla Turgut
University of Central Florida, Department of Computer Science, AIThings Research Lab
sharare.zehtabian@knights.ucf.edu, siavash.khodadadeh@knights.ucf.edu, Ladislau.Boloni@ucf.edu, Damla.Turgut@ucf.edu

Abstract

In the near future, augmented/virtual reality (AR/VR) characters might be used for tasks in the home that are currently performed through cell phones or laptop computers - ranging from checking the weather or news to performing banking, visiting a doctor, or going to school. Instead of a keyboard or touch screen interface, the user will interact with a virtual or real person, visualized life-size, with high quality through large screens or AR/VR devices. User satisfaction for such applications depends on delivering high-quality content with minimum latency. In this poster we describe a technique where we predict the user's future requests, use the prediction to prefetch the data from the network, cache it on a local device and show it to the user at the right time with minimum latency and maximum quality.

We describe a deep learning technique to predict the AR/VR experiences that the users are most likely to access at a specific time of the day and develop several different caching techniques. We rely on real-world smart home datasets, augmented with synthetic data created to match the essential attributes of the real-world data. We evaluate the proposed prediction methods and calculate the user's experience scores in terms of caching costs and user satisfaction. Finally, we compare our results with other baselines such as random caching, caching everything, and oracle. We found that our predictive approaches outperform the baselines, the difference being especially significant for the high-quality format deliveries.

Motivation

- AR/VR experiences require the delivery of **large** amounts of **high quality** data with **low latency**.
- For example: 3D weather or traffic report, major sport event.

- We consider the AR/VR for **daily** use within a household environment.
- The experience quality is limited by the:
  1. Capabilities of the devices through which it is delivered.
  2. Signal limitations such as network or bandwidth limitations.

Goals:

- **Predicting** what experience will the **user request** and in which **time frame** of the day,
- Designing **intelligent controller** for making decisions about the **forms and quality** of the contents.

User Modeling

- Learning models of the user behavior to predict the experiences the user will request.
- Scarcity of training data is a major challenge.

  Proposed solution:
  Creating synthetic data from real-world and simulated datasets of human daily activities.
  How?
  We probabilistically associate certain experiences with activities that are present in the dataset using **common-sense associations.**
  Example: weather forecast might be more likely to be accessed before leaving home.

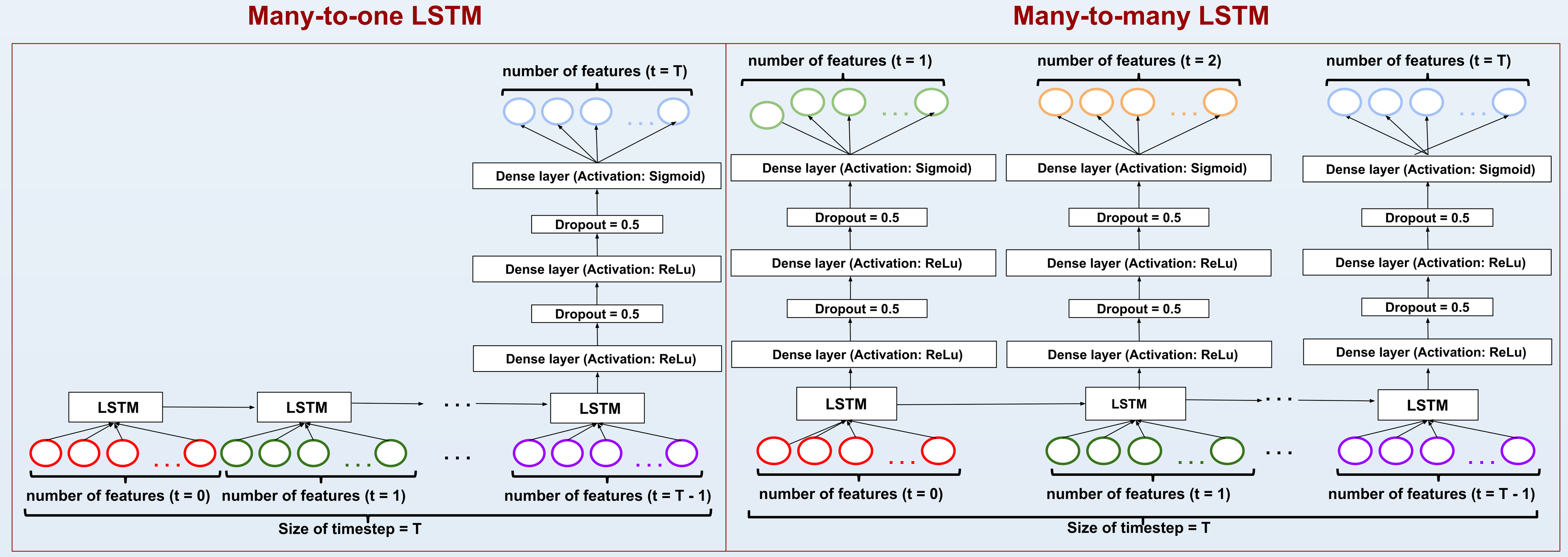
Task (Real-world Dataset 1)	Task (Real-world Dataset 2)	Task (Simulated Datasets 1 and 2)	Corresponding Request
Shave, Brush teeth, Get a drink	R1 wake, R2 wake	Other (50\% of the times), Leisure(60\% of the times)	Summary of news
Get dressed, Prepare for leaving (30% of the times)	Breakfast (70% of the times), Leave home (30% of the times)	Other (15\% of the times), Work (30\% of the times)	Weather repor
Prepare for leaving (50% of the times)	Leave home (50% of the times)	Other (10\% of the times), Work (50\% of the times)	Traffic report
Prepare for leaving (20% of the times)	Leave home (20% of the times)	Other (5\% of the times), Work (20\% of the times)	Parking status
Prepare brunch, Prepare dinner	Breakfast (30% of the times)	Other (20\% of the times), Leisure(40\% of the times)	Recipe

Intelligent Controller for Predictive Caching in AR/VR

We implement **three** different approaches to predict users' future requests:

1. **Probability-based**: is based on the probability of a specific request in a specific time interval.
2. **LSTM-based**: can not only process single data point, but also the entire sequences of data.
3. **Majority vote-based**: uses 15 different LSTM-based classifiers (by altering the hyperparameters).

Many-to-one LSTM **Many-to-many LSTM**



Hyperparameters for Majority Voting	Values
learning rate	0.001, 0.01
number of epochs	225, 300, 500, 1000
number of dense layers	1, 2, 3
regularization method	droptout(0.0, 0.2, 0.5, 0.8), l1 and l2

Baselines

1. **Oracle:** is considered as the best caching algorithm with 100% prediction accuracy.
2. **Cache everything:** caches every possible experience
3. **Random caching:** caches a randomly chosen request

Experimental Validation

Performance Metrics: Prediction accuracy & Final score

$$score(e_i) = d_d^{d(e_i)} \cdot d_f(f(e_i)) \cdot max\_score$$

delay discount format discount

$$final\_score = \alpha \cdot score - \beta \cdot cost$$

	Real Dataset 1			Real Dataset 2		
Caching Algorithm	4k video	HD video (1080p)	3D animation	4k video	HD video (1080p)	3D animation
Oracle	0.89	0.60	0.93	0.97	0.62	0.98
Cache everythin	0.00	0.31	0.39	0.00	0.31	0.39
Random	0.06	0.32	0.39	0.04	0.30	0.40
Probabilty based	0.40	0.44	0.62	0.62	0.51	0.76
LSTM based	**0.43**	**0.45**	**0.64**	0.60	0.50	0.75
Majority voting	0.37	**0.45**	0.63	**0.71**	**0.54**	**0.80**

	Simulated Dataset 1			Simulated Dataset 2		
Caching Algorithm	4k video	HD video (1080p)	3D animation	4k video	HD video (1080p)	3D animation
Oracle	0.93	0.61	0.96	0.94	0.61	0.96
Cache everythin	0.00	0.31	0.39	0.00	0.31	0.39
Random	0.00	0.45	0.65	0.45	0.45	0.66
Probabilty based	0.83	0.57	0.89	0.78	0.56	0.89
LSTM based	**0.90**	**0.60**	0.94	0.85	0.58	0.91
Majority voting	**0.90**	0.59	**0.93**	**0.87**	**0.59**	**0.92**

- LSTM-based and majority vote-based approaches outperform other approaches and provide maximum quality of delivery.
- For the lowest quality of delivery, we do not see that much difference between approaches.

Conclusions & Future Work

- We proposed an approach to perform a local caching of AR/VR experiences for a household scenario.
- We compared 3 different approaches: probability-based, LSTM-based and majority vote-based.
- Future work may include i) creating dataset of users' daily requests by considering the privacy and security of the users and ii) designing a collaborative learning prediction system among users.

Acknowledgements

The support for this work was provided by the **National Science Foundation** under Award **No. 1800961**. Any opinions, findings, and conclusions and recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

1. S. Zehtabian , S. Khodadadeh, L. Bölöni, D. Turgut. Intelligent controller for predictive caching in AR/VR: modeling and analysis of daily living in-home scenarios, Journal submission under revision, June 2020.
2. S. Zehtabian , M. Razghandi, L. Bölöni, and D. Turgut. Predictive Caching for AR/VR Experiences in a Household Scenario, In Proc. of IEEE Int'l Conf. on Computing, Networking, and Communications (ICNC 2020), February 2020.