Deep Auto-Encoding for Context-Aware Inference of Preferred Items' Categories

Moshe Unger, Bracha Shapira, Lior Rokach, Ariel Bar
Ben-Gurion University of the Negev
and Telekom Innovation Laboratories at BGU
Beer Sheva, Israel
mosheun@post.bgu.ac.il, {bshapira, liorrk, arielba}@bgu.ac.il

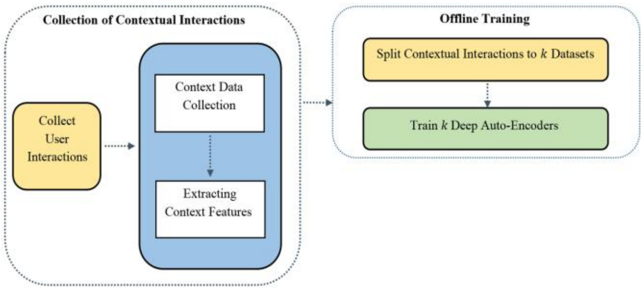
ABSTRACT
Context-aware systems enable the sensing and analysis of user context in order to provide personalized services to users. We observed that it is possible to automatically learn contextual factors and behavioral patterns when users interact with the system. We later utilize the learned patterns to infer contextual user interests within a recommender system. We present a novel context-aware model for detecting users' preferred items' categories using an unsupervised deep learning technique applied to mobile sensor data. We train an auto-encoder for each item genre, using contextual data that was obtained when users interacted with the system. Given new contextual sensor data from a user, the discovered patterns from each auto-encoder are used to predict the category of items that should be recommended to the user in the given context. In order to collect rich contextual data, we conducted an extensive field study over a period of four weeks with a group of ninety users. The analysis reveals significant insights regarding the inference of different granularity levels of categories that are available within the data.

Keywords
Recommender Systems; Deep Learning; Auto-encoder; Context; Mobile

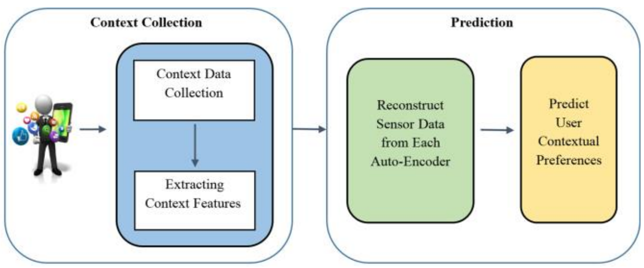
1. INTRODUCTION
Context plays an important role in determining the relevance of a service provided by an application to the user's needs. A system is considered context-aware if it can extract, interpret, and use context information and adapt its functionality to the users' immediate context [2]. Obtaining explicit contexts is a resource demanding task, since it requires either inputs from users, the knowledge of a domain expert, or collecting labeled contextual information. However, it is possible to use available users' ratings in order to learn implicit behavior patterns from available raw data [1,4,5]. In this paper, we build on Hull et al. (1997) who defined context as the user situation. When dealing with context, three entities can be distinguished: places (rooms, buildings etc.), people (individuals, groups), and things (physical objects, computer components etc.). In order to improve context prediction, it is necessary to discard indiscriminative or highly correlated features in order to avoid the curse of dimensionality [1]. In addition, when dealing with high dimensional data, it is important to consider dependencies between characteristics, in order to avoid a large number of model parameters. Baltrunas et al. [1] suggested a context based splitting approach in which ratings of certain items were split according to the value of an item-dependent contextual condition. Although this technique reveals the best contexts for each item, it is limited to a single and binary context and thus cannot model relations between several contexts. Moreover, while Baltrunas suggest to split the context for each item, we learn all context patterns that related to the same category of items. We suggest a novel approach for modeling and inferring users' context utilizing high dimensional data. We sense the user's rich contextual feature space (e.g., Wi-Fi networks, accelerometers, light, microphones, etc.) from the user's mobile phone and apply an unsupervised deep learning technique that extracts the most important features and discovers significant correlations between them. We rely on the available users' feedbacks in order to detect different behavioral patterns from the raw sensor data. Specifically, we split the data according to similar genres (or categories) of items that have been rated by users in a recommender system (RS). Then we apply auto-encoding on the divided data for each genre (e.g., auto-encoding for the 'food' genre, auto-encoding for the 'nightlife spot' genre, etc.) in order to discover different context patterns. Exploiting the implicit correlations between environmental features that affect user preferences, can be used to model the dynamic context of a user. Thus, we aim to obtain unsupervised contexts from the deep layers in order to determine the type of items relevant to the user's current context.

The major contributions of this paper include the following: first, we show how to infer contextual user preferences and availability employing the user's current unsupervised context and context models that were learned from past users' interactions with a RS. We suggest to split the rated data according to different genres/categories of items (e.g., "food," "nightlife spot") and learn a different deep model for each category by its contextual data. The models represent implicit contextual situations related to each category. Given new raw sensor data, we predict the category of an item that reflects the user's current context. The categories may be defined according to the level of granularity required by the target system's functionality and the available information about the categories. Second, we use an auto-encoders for modeling users' contexts from the data collected from mobile device sensors. We demonstrate our finding with data collected from real users during an extensive user study.

2. METHOD
Our method infers contextual preferences of users in terms of the users' availability for receiving recommendations and their preferred categories of items. The method consists of two phases, as presented in Figure 1: The first is the training phase in which we apply auto-encoding on contextual data collected from users' interactions within a recommender system. We build several deep neural networks for different splits of the data according to the items' categories (Figure 1a); The second phase is the prediction phase, where we use new sensor data currently recorded from the users' mobile phone. We utilize the learned deep models to reconstruct the input data and select the network that best fits the data with minimal error. We can then predict the current preferred user's category according to the selected contextual model (Figure 1b).



(a) Training Phase



(b) Prediction Phase

Figure 1: Method Overview

3. EVALUATION

3.1 Field Experiment and Data Collection

We aim to infer availability for receiving recommendations ("busy") and preferred categories of items ("food," "nightlife spot") collected from RS in order to provide meaning and further explanation regarding the unsupervised contexts. We evaluated our method on data that was collected from mobile device sensors in relation to a recommender system that provided recommendations of points of interest (POIs) obtained from Foursquare[1] API and received users' feedback about the provided recommendations. We developed an Android application which monitors the user's sensors and recommends popular POIs nearby.

90 students between the ages of 20-45 (53 male and 37 female) participated in the experiment. The overall experiment was conducted for a period of a month. Overall, the system collected 21,397 instances of positive user feedback (11,051 regrading food POIs, 812 regarding nightlife POIs and 9,534 reported "busy").

In order to acquire data from a variety of sensors, we used Android APIs and the Funf [3] application and collected nine sensors: accelerometer, Wi-Fi, battery, light, orientation, magnetic field, gravity, audio level, and location. Additional information was derived as well, such as day of the week, time of day, weather conditions, activity recognition, screen-log, call-log, and traffic statistics of certain applications installed on the user's mobile device (i.e., Facebook, WhatsApp, etc.). Raw data collected by the sensors and the additional information listed above were aggregated, analyzed, and engineered to generate 247 features.

We aimed to train a model that would be able to classify a test set of categories at different granularity levels. Specifically, we trained three auto-encoders that distinguished between three different contexts: food, nightlife spot, and busy. Second, we distinguished between seven contextual categories including: restaurant, vegetarian, fast food, desserts, café, nightlife spot, and busy. We applied several prediction models, using the time-based splitting method, on the test data. Table 1 presents results for prediction of the two settings. As observed, our model significantly outperforms all of the tested classification algorithms by at least 45% in terms of accuracy and by 23% in terms of AUC. This phenomenon can be explained by the fact that the auto-encoder can better represent high dimensional data and correlations between features using all available features. We can also notice that the tested classifiers obtained much lower results in the prediction accuracy of seven categories. This can be explained by the fact that the detailed items' categories are conceptually more similar to each other, compared to categories in previous setting.

[1] https://developer.foursquare.com/

Table 1. Prediction of High Level Categories

Prediction Model	Prediction of Three Categories		Prediction of Seven Categories	
	Accuracy	AUC	Accuracy	AUC
Auto-Encoding	**0.971**	**0.927**	**0.884**	**0.918**
Random Forest	0.657	0.739	0.484	0.685
SVM	0.655	0.671	0.5	0.617
C4.5	0.668	0.751	0.51	0.289

4. CONCLUSION

In this paper we presented a novel approach for inferring contextual user preferences by applying auto-encoding on sensor data. Our solution relies on the identification and usage of positive feedbacks acquired in contextual situations. In order to evaluate our suggested model, we conducted an extensive user study over a period of four weeks with a developed application which displays POI (point of interest) recommendations to users. The experimental results show that we were successfully able to predict preferred items' categories at different granularity levels. In all settings the auto-encoding approach was superior to the traditional state-of-the-art classification methods in terms of accuracy and AUC. The results indicate that auto-encoder is the most effective method tested for modeling contextual patterns when dealing with high dimensional data.

5. REFERENCES

[1] Baltrunas, Linas, and Francesco Ricci. "Context-based splitting of item ratings in collaborative filtering." Proceedings of the third ACM conference on Recommender systems. ACM, 2009.

[2] Byun and Chverst B. Sarwar, G. Karypis, J. Konstan, and J. Riedl. "Analysis of Recommendation Algorithms for ECommerce". Electronic Commerce Proceedings of the 2nd ACM conference on Electronic commerce, 158 – 167, 2000.

[3] Nadav Aharony, et al. "Social fMRI: Investigating and shaping social mechanisms in the real world." Pervasive and Mobile Computing 7.6 (2011): 643-659.

[4] Guo, Lei, et al. "Learning to recommend with social contextual information from implicit feedback." Soft Computing 19.5 (2015): 1351-1362.

[5] Wang, Hao, Naiyan Wang, and Dit-Yan Yeung. "Collaborative deep learning for recommender systems." Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2015.