A Deep Learning Framework for Product Recommendation

Contributors:

Vishwanath Prudhvi
Aditya Dutta
Chandramauli Chaudhuri
Aniket Koyande
Yuvaneet Bhaker

January 2017

Agenda

1. Context
2. Problem statement
3. Overview of the data
4. Motivation and candidate approaches
5. Solution approach
6. Results and technical specifications
7. Next steps

Context

Driving customer centricity

- Across customer-facing industries like retail, insurance, banking etc., customer lifetime value is the most critical factor driving revenues & long term profits
- It is thus vital for businesses to maximize customer loyalty and lifetime value by ensuring a personalized experience across their journey
- Hence, Next Best Product/Action frameworks are becoming increasingly more important, keeping in view the shift from product-centricity towards customer-centricity

Data availability

- In high touch-point industries like retail, asset management and banking the data available is rich and granular down to individual customer ID level
- External customer-related data is easily accessible, which can substantially improve prediction accuracy
- This can be used to mine patterns across purchase/usage behavior and is suitable for building deep learning frameworks

Next Best Product

Problem statement

One of the largest retail banks in Europe wants to predict the **incremental product usage** at individual customer level. The available dataset is demographic and historical product usage information. The evaluation metric to be used is MAP@7.

Usage and *Incremental Usage* are defined as follows:

$$\textit{Usage}(t, i, j) = \begin{cases} 1 & \text{if customer } i \text{ has used product } j \text{ during time period } t \\ 0 & \text{otherwise} \end{cases}$$

$$\textit{Incremental Usage}(t, i, j) = \begin{cases} 1 & \text{if } \textit{Usage}(t-1, i, j) = 0 \text{ and } \textit{Usage}(t, i, j) = 1 \\ 0 & \text{otherwise} \end{cases}$$

Driving key business decisions

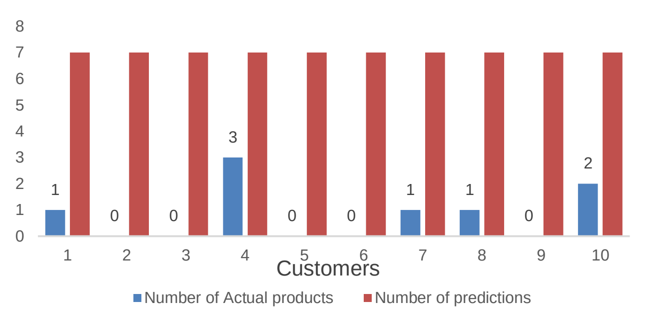
With an effective recommendation system, the bank can better meet the individual needs, personalize customer experience and ensure greater customer satisfaction

Evaluation metric: MAP@k (Mean Average Precision)

What is MAP@k ?

***MAP@k** for a set of customers represents the mean across the average precision scores of each individual customer*

MAP@k may return low scores



Precision

- % of selected items that are correct
- $Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$

Relevance

- Selected item is relevant
- Rel(i) = 1 if relevant product, else 0

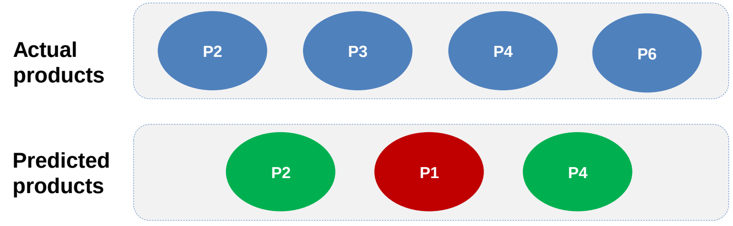
Average Precision at k

- $\mathrm{AP@k} = \frac{\sum_{i=1}^{k}(\mathrm{P(i)} * \mathrm{rel(i)})}{\mathrm{k}}$
- $\mathrm{k} =$ total number of relevant items

Mean Average Precision@k

- $\sum_{j=1}^{M} \frac{\mathrm{AP@k(j)}}{\mathrm{M}}$
- $\mathrm{M} =$ total number of customers

Illustrative example for estimating MAP@3

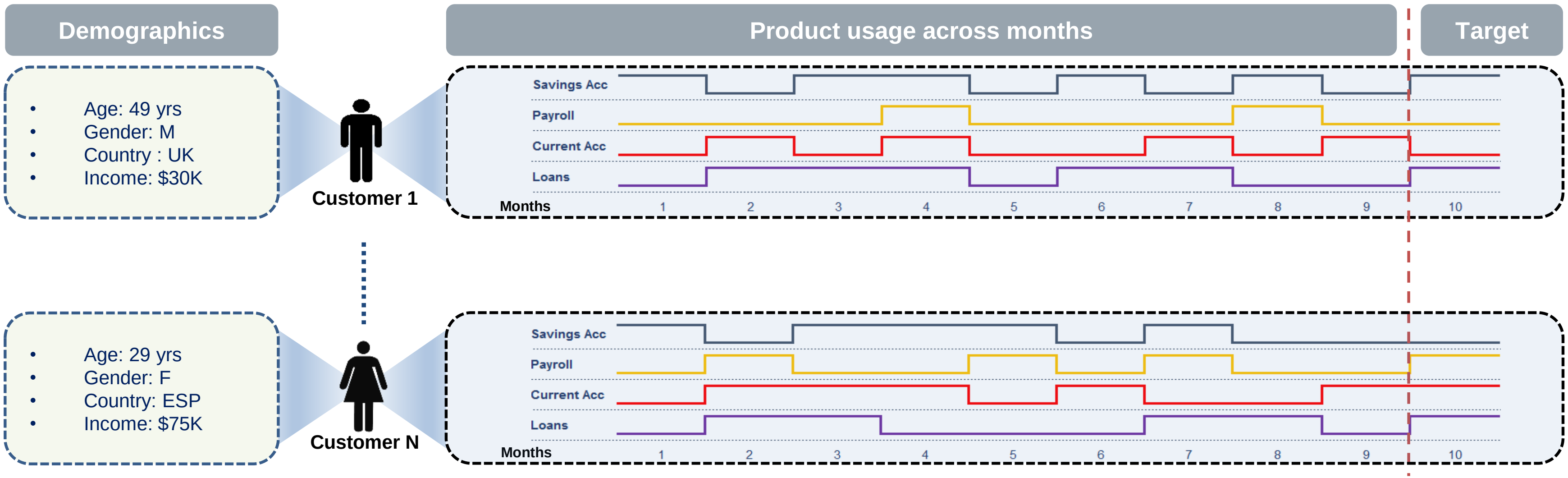


	1	2	3
Precision at i	1/1	1/2	2/3
Relevance at i	1	0	1

AP@3 = (1.00*1 + 0.50*0 + 0.67*1)/3 = 0.57

MAP@3 is obtained by averaging AP@3 of each customer across the entire population

Overview of data



- 1 million unique retail banking customers (total sample of approx. 14 million)
- 1.5 years of data aggregated at monthly level
- Data on usage of 24 distinct retail banking products like credit card, loans etc.
- 24 demographic variables including age, gender, income etc.

Motivation and candidate approaches

Traditional approaches

Approach	Limitations
Using business rules	Not good enough to build an exhaustive customer-specific model, can be used in post-processing
Decision trees	Do not perform well if the number of classes is very high as typically observed in BFSI domain
Collaborative Filtering	Takes into account just the relation between user and item, ignores context information

- None of the traditional approaches account for sequence information (product usage sequence in our case)

Motivation behind using deep learning architecture

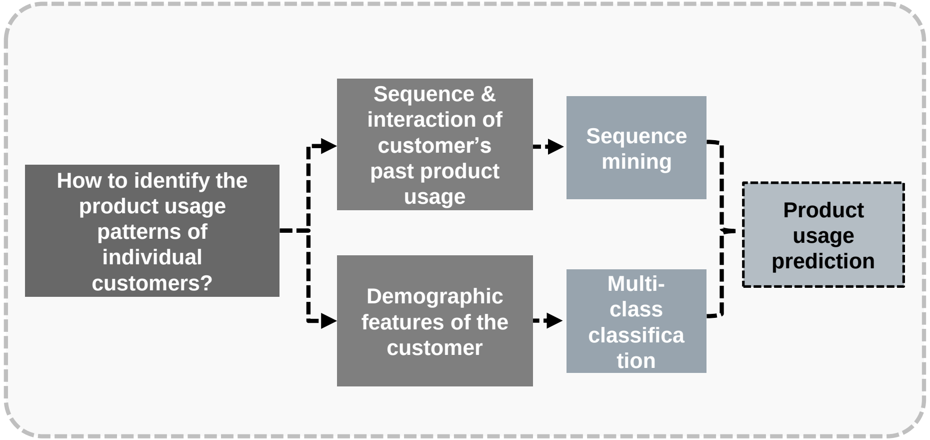
- Provides a comprehensive framework that can incorporate sequence information and demographic data
- Takes into consideration interactions between products
- Intrinsic nature of product sequences can be captured using RNNs
- Prior experiments on Amazon’s product ratings data at Customer Genomics had shown promising results

Solution approach: Overview

1 Identify product usage patterns

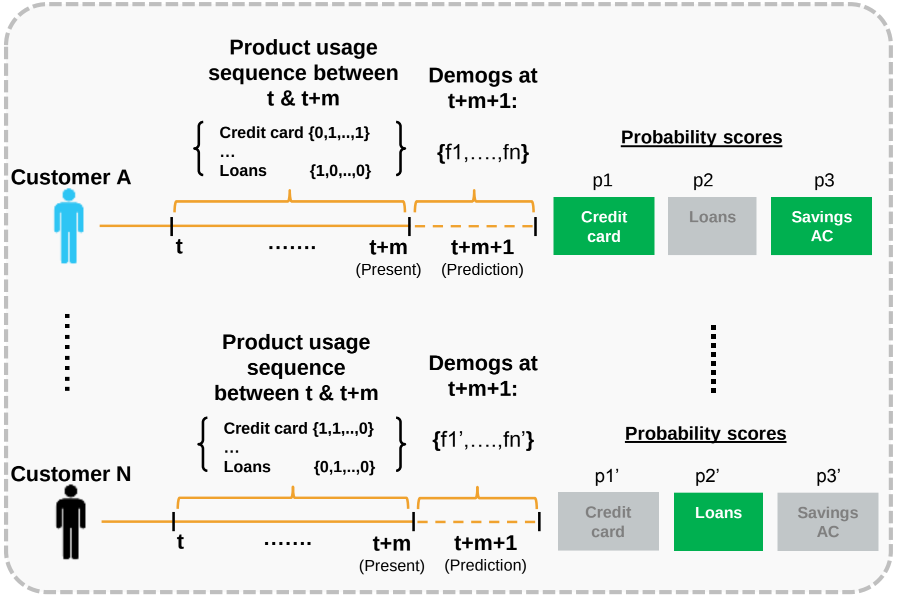
To predict the next product usages for retail banking customers, it is important to identify their product usage patterns from the past data. There are 2 aspects to understanding this behavior:

- **Sequence and interaction of customer's past products usages**
- **Demographic features of the customer**



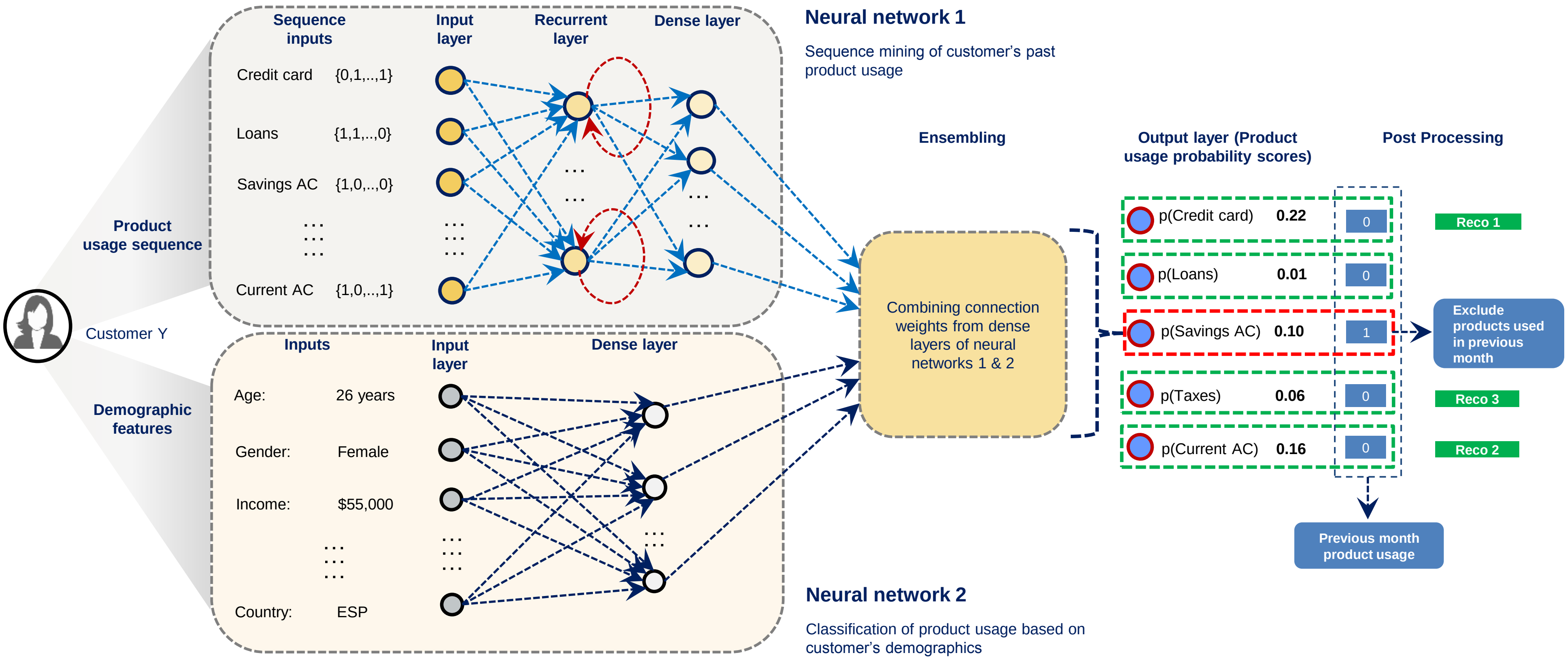
2 Predict future product usage

Predict the probability of usage of each product at a future time-period t+m+1 based on product usage sequence between t to t+m and demographic features at t+m+1



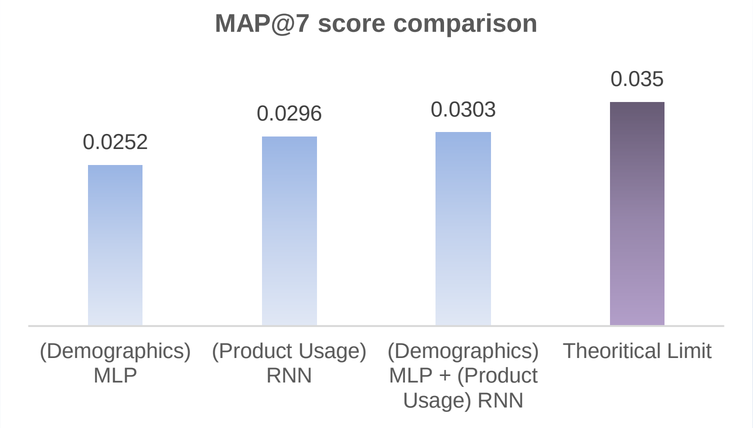
*** Index: 1 - usage, 0 – non-usage**

Solution approach: Deep learning architecture



Results and technical specifications

Evolution of product usage recommender



Technical specifications

- Model run time: 30-40 minutes (5-10 epochs)
- 32 GB RAM
- 8 Cores
- Python libraries: Keras-1.1.2 with Theano-0.7.0 backend

- Evaluation metric: MAP@7
- Out-of-time validation to ensure accuracy and avoid overfitting
- Drop-out layers to avoid overfitting
- Achieved a final MAP@7 score of 0.0303 which is 85% of the maximum theoretical score of 0.0350

Next Steps

Next steps

- Generalized framework which can work well even on extremely low frequency product usages (<5%)
- Experimentation on alternative loss functions that can be used in this context
- Ready-to-implement Next Best Product module that can be installed across multiple BFSI clients