The Customer Revisit Prediction Using Macroscale Mobility Information

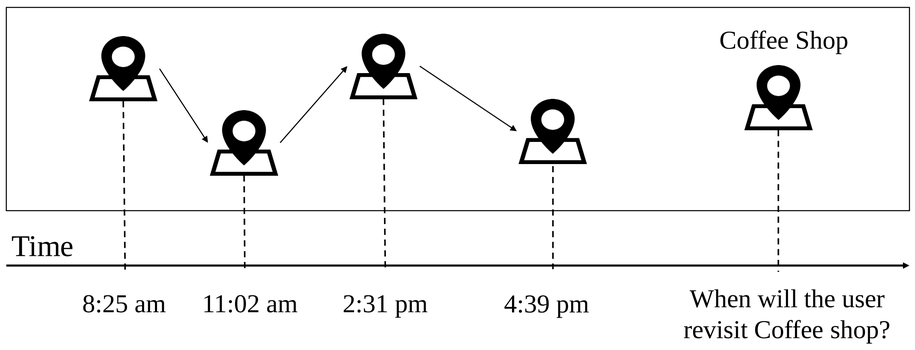
Beomyoung Kim (dglidgli@kaist.ac.kr)

Department of Industiral and Systems Engineering , KAIST

Introduction

Motivation

There is a research that successfully predicts the revisit of customers by using only movement data collected in a specific offline store through wireless fingerprinting technology.[1] Companies that have macroscale mobility information, such as Loplat X, also want application in a variety of areas, such as target marketing and customer segmentation through revisit prediction.



Purpose

1) To present model for predicting revisit time for specific place categories or brand rather than a single store.

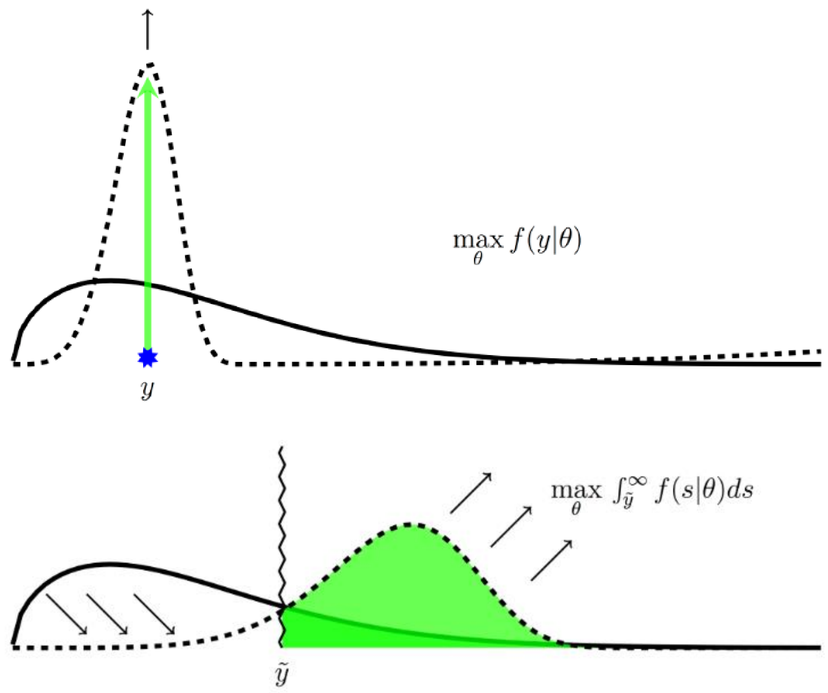
2) To solve the data scarcity problem, which is not known whether the customer visited after that time when using data up to a certain point for training model.

3) To define features and embedding factors related to revisit in macroscale mobility information and to extend the WTTE-RNN methodology.[2]

Method

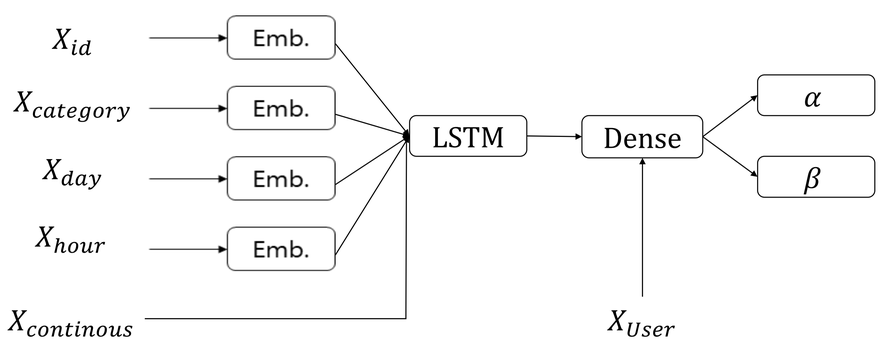
Loss Function

When censored observations, maximize the probability over the right of that point. When uncensored observations, maximize the density function of that point.[2]



$$\sum_{n=1}^{N}\sum_{t=0}^{T_n} u_t^n * \log[\Pr(Y_t^n = y_t^n|x_{0:t}^n)] + (1-u_t^n) * \log(\Pr(Y_t^n > y_t^n|x_{0:t}^n)]$$

Model



Assume the function of revisit time follows Weibull distribution.

Experiment

Features

Embedding Features

1) Category 2) Day of the week 3) Hour 4) Customer Id

Continuous Feature

1) Time until next any log

Aggregated Features

* Also used in Cox Model

1) The number of interest category over the entire period

2) The number of all logs 3) The number of days with logs

4) The number of all logs divided by the days with logs

5) Time from last visit of interest category

6) The number of weekend over the entire period

7) Average interval between interest category

8) Average interval between all logs

Results of "Coffee Shop"

	Cox	**WTTE-RMM**
C-index	0.698	**0.726**
RMSE	205.24	**199.49**
Non-returning recall	**0.488**	0.438
Non-returning F1	0.543	**0.555**

Ours has less error in actual revisit time and can classify for customers who do not revisit compared with cox which is one of the existing survival analysis models.

Reference

[1] Sundong Kim and Jae-Gil Lee, "Utilizing in-store Sensors for Revisit Prediction," IEEE international Conference on Data Mining, pp. 217-226, 2018.

[2] Egil Martinsson, "WTTE-RNN : Weilbull Time to Event Recurrent Neural Network," Chalmers University of Technology, Gothenburg, Sweden, 2016.