User response prediction in mobile advertising
Faustine Bousquet, Khanh Duong, Christian Lavergne, Sophie Lèbre, Anastasia Lieva

To cite this version:
Faustine Bousquet, Khanh Duong, Christian Lavergne, Sophie Lèbre, Anastasia Lieva. User response prediction in mobile advertising. ECML PKDD, Sep 2018, Dublin, Ireland. hal-02014821

HAL Id: hal-02014821
https://hal.archives-ouvertes.fr/hal-02014821
Submitted on 11 Feb 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Aims and objectives

Our aim: targeting the right person, at the right place and time with the most relevant ad

- Clustering of mobile campaigns
  - Each campaign has its own KPI to optimize. It can be a performance objective (lots of clicks), a branding strategy (lots of impressions) or more complicated goals which are not easy to handle through bid requests. Then, first part of the thesis focuses on obtaining clusters of campaigns.
  - Click prediction models for each cluster

Mathematical approach

Model definition
Impressions and clicks are aggregated by hour. We calculate the corresponding CTR or number of impressions for each time slot. Observations can be described as:

For all

- \( c = 1, \ldots, C \) campaigns,
- \( j = 0, \ldots, J \) days of campaign \( C \),
- \( h = 1, \ldots, H \) time slots,
- \( t = 1, \ldots, T_j \) repetitions of time slot \( h \) during day \( j \),

where \( \mu \) is a constant, \( \beta_k \) the time slot effect, \( \beta_{k}^{d} \) the day of week effect (assuming the following constraints: \( \beta_{k}^{d} = 0 \) for identifiability) and \( \epsilon_j \) a gaussian error .

Mixture model

We assume that there are \( C \) campaigns which are part of \( K \) groups:

\[
Z_C = \begin{cases} 
1 & \text{if campaign } c \text{ belongs to cluster } k \\
0 & \text{otherwise}
\end{cases}
\]

The mixture model can then be written like:

\[
f(y_{i,j} | \beta, \sigma^2) = \sum_{k=1}^{K} \lambda_k f(y_{i,j} | \beta_k, \sigma_{k}^2)
\]

where \( f(y_{i,j} | \beta, \sigma^2) = \Pi_{c=1}^{C} \Pi_{k=1}^{K} f(y_{i,j} | \beta_k, \sigma_{k}^2) \) and \( P(Z_C = 1) = \lambda_k \), the probability that \( Y \) belongs to \( k \).

We used a classical Expectation-Maximization (EM) [1] algorithm to estimate this mixture model.

Criterias for choosing number of cluster

- Bayesian Information Criterion (BIC) \( BIC(K) = -2 \log(L(Y, \beta, \sigma^2)) + \phi \text{parameters} \times \log N \) where \( N \) is the number of observations.
- Integrated Classification Likelihood [2] (ICL) which penalizes the complete log likelihood:

\[
ICL(K) = 2 \sum_{c=1}^{C} \sum_{k=1}^{K} \lambda_k \log \lambda_k \text{ where } \lambda_k = P(Z_C = 1 | Y)
\]

References