

# スマートフォンのセンサ情報を用いた屋内来訪検知技術

水上 ひろき<sup>1</sup> 尾日向 洋皓<sup>2</sup> 横山 潤<sup>2</sup>

**概要：**本研究では、スマートフォンのセンサ情報を持ちいた、屋内の特定のエリアへの来訪を計測する技術に関して技術的な検証結果を紹介する。屋内における来訪計測では多くの場合来訪計測の対象地点に Beacon などの特定の機器を設置することが多いが、この取組みは機器の設置を行わずに屋内での来訪測定精度の改善を達成した。

## Indoor Visit Detection Technics with Sensors of Smartphones and Machine Learning

Hiroki Mizukami<sup>1</sup> Hiroaki Obinata<sup>2</sup> Jun Yokoyama<sup>2</sup>

### 1. 言葉と記号の定義

本研究では問題の背景としてインターネット広告の話題を取り扱う。その中で広告を配信したい商業主体のことを広告主 (advertiser) と呼び、広告の掲載面を取り扱う広告媒体のことをメディア (media)、広告配信対象となる一般聴衆のことをオーディエンス (audience) と呼ぶ。また、スマートフォンのあるアプリケーションの利用者をユーザ (user) と呼ぶが、とくに広告配信の対象として取り扱う場合はこれをオーディエンスと呼ぶこともある。

### 2. 背景

#### 2.1 インターネット広告とセンサデータ

これまで、インターネット広告の影響量の指標及び KPI として様々な値が用いられてきた。例えばオーディエンスの振る舞いに関する指標としては CTR (クリック確率) や CVR (成約確率) などがあり、費用対効果を評価する指標として CPA (単位成約あたりコスト)、CPC (クリックあたりコスト) 等がある。

新しい評価指標が現れるに伴いその正確な計測方法がインターネット広告の大きな課題となってきた。例としては広告バナーのクリックが広告媒体と広告事業者のみで計測出来るのに対して、成約を観測するには広告主の協力が欠

かせないものとなるなど、一般にオーディエンスの行動をより詳細に調べようとするほど、その効果の計測には複雑なしくみが必要になることが多いのである。

この評価指標の模索の延長線上に、実店舗への来訪誘導効果がある。しかしインターネット広告において、広告接触者の実店舗への来訪を追跡することはとても困難を極める。したがって小売のような事業者の企業は折込チラシと会員カードの利用ログを組み合わせたなど、インターネットとは切り離された広告媒体のみでその効果の計測を行ってきた。

一方で近年では技術の発展に伴いスマートフォンを用いて加速度や GPS などの様々なデータが取得出来るようになった。とくに GPS 信号などの位置情報の提供は、道案内をするとき、天気を調べるときなどもはや我々の生活には欠かせないものとなっている。このことを適切に用いればショッピングモールなどの大型施設に限り来訪の計測が期待されているが、高層階の特定のフロアなどの高さを区別する必要がある場合や、コンビニエンスストアなどの小さな店舗において、GPS による計測では精度が不十分であることがわかっている。

#### 2.2 センサデータ活用に関する課題とアプローチ

スマートフォンのセンサデータをマーケティングの現場に活用するには大きく 2 つの課題が考えられる。

一点目はそのデータの解釈に関する課題である。取得さ

<sup>1</sup> 株式会社サイバーエージェント 技術本部 秋葉原ラボ (※2017年3月までアドテク本部)

<sup>2</sup> 株式会社サイバーエージェント アドテク本部 AirTrack

れる未加工データは機械的なもので、人間が理解して知見を獲得するためには何らかの加工や分析が必要になる。短絡的には機械学習により直接クリックなどの成果の予測に用いる事ができる [Juan 11] が、「オーディエンスを理解したうえでより良い訴求方法や施策を考える」といった場合にはマーケティング上の属性（性別や趣味嗜好，行動特性など）を抽出することのほうが成果を直接予測するより重要だと考えられる。その意味で本研究は屋内来訪という点に着目したマーケティング変数の獲得を試みるものである。また予測などのタスクにおいてもマーケティング上の属性を用いることで予測の精度が上がる例も報告されている [新美 17]。

二点目はデータの鮮度である。スマートフォンのセンサデータのように絶え間なく発生し続けるデータをストリーミングデータというのが [Tyler 15]，このようなデータは発生した瞬間の価値が最も高い事が多い。一方で，生ログを加工したような中間データは日次のバッチ処理などで作成されること多い。たとえば日次バッチ処理の場合ログの発生と活用出来るまでの時間差が最大で24時間発生し，ここに大きな価値の損失が生まれるのである。これは特定の商品の検索のような行動ログを例に取れば，オーディエンスが該当の商品を購入したあとと購入する前ではそのデータの価値に雲泥の差があることが理解しやすいだろう。同様に今回のような来訪ログに関しても「オーディエンスがその場を離れる前」がデータとしての価値が最も高い瞬間となる。今回の検証における実装は先に上げたような価値の損失を防ぐために，予測と学習共にオンラインのロジックを採用することでリアルタイムな予測を可能にしているが，この論文の中では利用データとモデルに注目するために具体的なアーキテクチャについては触れないこととする。

### 3. 使用モデル

#### 3.1 データについて

与えられたセンサのデータからオーディエンスが来訪したエリアを判定するために教師ありの多値分類器を構築することを目指す。特徴量としては端末から取得できるGPSやWiFiなどデータを用いる。以下に学習データの例を示す。

```
{
  "label": 1,
  "feature": {
    "latitude": 35.699203,
    "longitude": 139.772154,
    "aa:aa:aa:aa:aa:aa": -84,
    "bb:bb:bb:bb:bb:bb": -76,
    "cc:cc:cc:cc:cc:cc": -73,
    ...
  }
}
```

ここで要素 label の値は， $i \in \{1, 2, \dots, m\}$  なる数で，教師信号としてはこの 1 of  $k$  表現である  $m$  次元ベクトルを用い，以後このベクトルを  $\mathbf{t} \in \{0, 1\}^m$  と表す。

#### 3.2 feature hashing

ここでは次元数が未知な特徴量を内積をおおよそ保ったまま固定長のベクトルに変換する feature hashing [Weinberger 09] について紹介する。

feature の要素は文字列のキーに対して実数値を対応づける構造となっている。いまキー全体の集合  $K = \{k_1, k_2, \dots\}$  とし，このそれぞれのキーに対応する値の列を  $\mathbf{v} = (v_{k_1}, v_{k_2}, \dots)$  と書くとする。さらに任意の  $k \in K$  に  $F = \{1, 2, \dots, f\}$  の元を対応させるハッシュ関数を  $h: K \rightarrow F$  と書くとする。これらを用いて

$$x_i = \sum_{k; h(k)=i} \xi(k)v_k$$

なる  $v$  に対応する  $f$  次元ベクトル  $\mathbf{x}_v = (x_1, x_2, \dots, x_f)$  を定義する。ここで  $\xi$  は任意の  $k \in K$  に対して 1 もしくは  $-1$  を割り当てる関数とする。このとき任意の2つデータ  $\mathbf{v}_1, \mathbf{v}_2$  について

$$E[\langle \mathbf{x}_{\mathbf{v}_1}, \mathbf{x}_{\mathbf{v}_2} \rangle] = \langle \mathbf{v}_1, \mathbf{v}_2 \rangle$$

となる事が知られている。ここで  $E[\cdot]$  により確率変数の期待値， $\langle \cdot, \cdot \rangle$  はベクトルや数列の内積とする。

今回はこのようにして得られた学習データの列  $(\mathbf{x}_1, \mathbf{t}_1), (\mathbf{x}_2, \mathbf{t}_2), \dots$  を教師データとして使用する。

#### 3.3 全結合多層ニューラルネットワーク

分類器としてはごく一般的な多層 ( $l$  層) ニューラルネットワークを用いる。中間層の活性化関数としては

$$a(x) = \max(x, 0)$$

を用いる。この関数  $a$  は **rectifier linear unit** もしくはその略称として **ReLU** と呼ばれており特定のタスクで良い成績を上げている [Glorot 11]。

入力層の値を  $h^{(0)} = \mathbf{x}$ ， $i$  層の重み行列を  $W_i$ ，バイアスを  $\mathbf{b}_i$  とするとき  $i = 1, 2, \dots, l-1$  なる中間層を以下で定める。

$$\mathbf{h}^{(i)} := a(W_i \mathbf{h}^{(i-1)} + \mathbf{b}_i)$$

また出力層に関しては

$$\text{softmax}(\mathbf{x})_i = \frac{e^{x_i}}{\sum_j e^{x_j}}$$

なる Softmax 関数を用いて

$$\mathbf{h}^{(l)} := \text{softmax}(W_l \mathbf{h}^{(l-1)} + \mathbf{b}_l)$$

により  $\mathbf{h}^{(l)}$  を定める。

| モデル           | 精度    | 再現率 |
|---------------|-------|-----|
| ロジスティック回帰     | 0.928 | 1.0 |
| 多層ニューラルネットワーク | 1.0   | 1.0 |

表 1 モデル毎の性能評価

### 3.4 学習と正則化

損失関数には交差エントロピー

$$\text{loss}(\mathbf{x}, \mathbf{t}) = - \sum_i t_i \log(h_i^{(l)})$$

を用い、逐次勾配降下法 [Bottou 98] によりこの損失を小さくする  $W_1, \dots, W_l, \mathbf{b}_1, \dots, \mathbf{b}_l$  を学習する。

また学習の際に汎化性能を高めるために Dropout [Srivastava 14] を用いている。Dropout は学習の際に各ノードを一定の確率  $0 < \alpha < 1$  で消した状態で学習を行う方法で、様々なデータで過学習を抑制することが知られている。また Dropout については理論的な解析も進んでおり、特定の状況下では確率的に  $L_2$  正則化に相当する事が知られている [Wager 13]。この  $\alpha$  はハイパーパラメタで今回は  $\alpha = \frac{1}{2}$  を用いている。

## 4. 実験

### 4.1 屋内における手法の比較

自社のオフィス内を 2 つの区画に分けてそれぞれでデータを取得した。分類器の評価のために、このデータを用いて今回の多層ニューラルネットワークのモデルとロジスティック回帰モデルの精度の比較を行った。2 つのモデルの精度と再現率を表 1 に示す。

この結果から多層ニューラルネットワークのほうが精度の意味で良いことがわかるが、これはニューラルネットワークが非線形な性質を分類できる事が理由と考えられる。

### 4.2 都内における大規模実験

測定対象とする屋内の商業施設などの 94 区画において学習用の合計 41189 行を取得し学習する実験をおこなった。対象となる区画はいずれも屋内で、特定のフロアや、会議室、テナントなど従来の GPS による来訪検知が不可能なものである。

feature hashing に用いるハイパーパラメタには  $f = 6656$ 、多層ニューラルネットワークの層の数は  $l = 4$ 、またそれぞれのノードの数は最初の層から順に 2048, 2048, 1024, 1024 とした。学習データのうち 19997 行は市街地の屋外など無関係の区画で取得したデータでこれを擬似負例として学習させた。

精度の評価について h 5-fold の交差検証を行った。計測対象区画毎に精度と再現率計算し、図 1 でそのヒストグラムを示す。精度、再現率とそれぞれで 0.9 以上の区画が 8 割以上を占める。

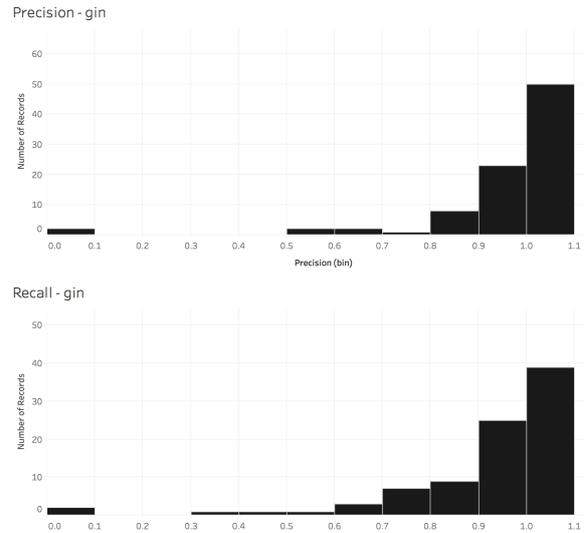


図 1 区画ごと精度と再現率のヒストグラム

## 5. 考察

### 5.1 アプリケーションにおける端末情報の活用と個人情報

アプリケーションが、加速度、GPS など端末のセンサ情報を利用する際は、必ず利用目的を明確に説明したうえでユーザの同意と共に利用されるべきである。

傍ら本検証の結果は、機械学習などのテクニックを用いることで、端末に搭載されセンサ本来の役割を超えた情報が得られる可能性も示唆している。たとえば、今回検証するスマートフォンのセンサデータが屋内の来訪判定に用いられる事は、専門知識を持たない者からすれば想像に難しいことかもしれない。しかしながら今回の場合もアプリケーションで用いられる場合は位置情報を推定するためになどのセンサデータを利用し、それがアプリケーションの機能改善に利用されるという点についてユーザに予め合意をとることが、健全な情報活用のためには重要だと考えられる。(多くの OS のアプリケーションにおいて、センサデータを利用する場合は位置情報利用に関する許諾が必要になっている)

### 5.2 まとめと今後の課題

今回のモデルにより、機械的なセンサデータの列から「特定区画への来訪」というユーザの性質を抽出することが可能になり、その性能も十分実用に耐えるものとなった。また、feature hashing により後日特徴量を追加した際にも追加の学習と予測を可能にしており、データの取得から活用までの時間差を短くすることを可能にした。

一方で分類できる地点のには限りがある。例えば今回採用したモデルの出力層は 1024 次元のベクトルだが、これ

はモデルでは 1024 地点以上の分類ができない事を意味している。この出力層の数は予め決める必要があるため、層の数やノード数などのモデルの規模と分類可能な区画の数の関係は今後も調査が必要となる。

また、今回採用したセンサ情報には同一区画においても、時間とともに変化する可能性があるものも含まれている。これは時間の経過とともに次第に分類精度が保証されなくなって行くことを意味する。モデルの鮮度を維持するために継続的に教師データを取得する仕組みが出来るかどうかの実用化に向けては重要な課題となる。

## 参考文献

- [1] 新美 潤一郎, 星野 崇宏: 顧客行動の多様性変数を利用した購買行動の予測. 人工知能学会論文誌 Vol. 32 (2017) No. 2
- [2] Juan, Y.C., Zhuang, Y., Chin, W.S.: 3 Idiots Approach for Display Advertising Challenge. *Internet and Network Economics*, pp. 254265. Springer, Heidelberg (2011)
- [3] Tyler Akidau: The world beyond batch: Streaming 101, <https://www.oreilly.com/ideas/the-world-beyond-batch-streaming-101> (2015)
- [4] K. Weinberger, A. Dasgupta, J. Langford, A. Smola, and J. Attenberg. Feature hashing for large scale multitask learning. In *ICML' 09*. (2009)
- [5] Glorot, X., Bordes, A., and Bengio, Y: Deep sparse rectifier neural networks. In *AISTATS' 2011*. (2011)
- [6] Lon Bottou: *Online Algorithms and Stochastic Approximations, Online Learning and Neural Networks*, Edited by David Saad, Cambridge University Press, Cambridge, UK. (1998)
- [7] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.: Dropout: a simple way to prevent neural networks from overfitting. *J. Machine Learning Res.* 15, 19291958 (2014)
- [8] S. Wager, S. Wang, and P. Liang.: Dropout training as adaptive regularization. In *Advances in Neural Information Processing Systems 26*, pages 351359, (2013)