2D-02

自動 ID 取得ツール対策としての多層ニューラルネットワーク

安藤 義裕
ヤフー株式会社
yoandou@yahoo-corp.jp

概要

インターネットサービスの ID を自動で取得するツールがネット上で販売されている. このようなツールは主にアフィリエイトで収入を得ているアフィリエイターと呼ばれる人たちを対象としている. アフィリエイターは, 取得した ID に付与された無料のメールアドレスを使い, ブログサービス事業者のサイトでブログを開設する. そして, 開設したブログにアフィリエイトリンクを掲載することで収入を得ている. メールアドレスが多ければ多いほど, 開設できるブログの数が多くなり, 効率的に収益に結びつけることができる. ID およびメールアドレスが大量に取得されると, メールアドレスを提供する事業者にとっては, 管理コストが増加することになる. また, ブログを提供する事業者にとっては, アフィリエイト目的のブログが増えることにより, サービス全体のイメージが悪化することになる. このような背景をもとに, 我々は ID を自動で取得するツールを検知し, ID の取得制限につなげられる手法の検討を行った. 我々は画像認識等で高い性能を発揮しているディープラーニングに着目し, ディープラーニングで使用される多層ニューラルネットワークを用いて自動 ID 取得ツールの検知を行う手法を開発した. この手法を用いることにより, 自動 ID 取得ツールが高い精度で検知されることを確認した.

1. はじめに

インターネット検索サイトで“自動 メールアドレス 取得ツール” と入力して検索を行うと, 大手インターネットポータルサイトが提供している ID を自動的に取得するツールが多数見つかる. これらのツールは主にアフィリエイト (成功報酬型広告) で収入を得るアフィリエイターと呼ばれる人たちを対象としている. アフィリエイターは ID に付随する無料のメールアドレスを利用して複数のブログを開設し, そこにアフィリエイトリンクを設置することで効率的に収入を得ている. 図 1 はアフィリエイトで収入を得る流れを図式化したものである [1].

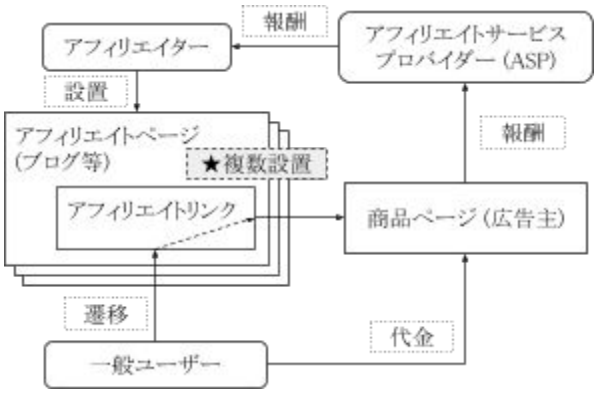

図 1: アフィリエイトの仕組み

複数の ID とメールアドレスを取得して複数のブログを開設すれば効率的に収益を得ることができる. このような行為は, ブログサービスを提供している事業者からすると, いわゆるスパムブログが増えることになり, 提供しているブログサービスのイメージ悪化につながる. また, メールアドレスを提供する事業者からすると, 管理コストの増加につながることになる. したがって, 自動で複数の ID を取得する行為は, インターネットサービスを提供する事業者にとって好ましいものとは言えない. 我々はこのような問題意識をもとに, 自動 ID 取得ツールを検知する方法を検討した. 我々は画像認識等で高い性能を発揮しているディープラーニングに着目し, そこで使用されている多層ニューラルネットワークで検証を行った. 我々が実装した多層ニューラルネットワークを利用し, 市販のツールで検証を行ったところ, 高い精度でツールの利用を検知できることを確認した.

2. 提案手法

2.1. 多層ニューラルネットワーク

多層ニューラルネットワークはニューラルネットワークの中間層を多層化したものである. 図 2 は本論文の多層ニューラルネットワークを図式化したものである.

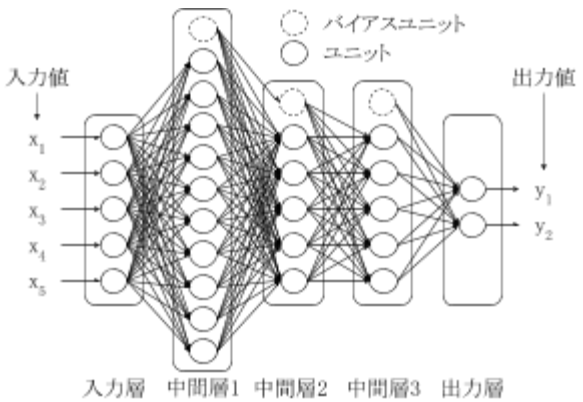

図 2: 提案手法の多層ニューラルネットワーク

我々のネットワークは入力層, 中間層 3 層, 出力層で構成され, ユニット数は入力層が 5, 中間層は入力層側

から 10, 5, 5 とした. 出力層は 2 で, それぞれ“ツールによる入力”, “人による入力”の 2 値分類をソフトマックス関数で実現するものとした. なお, 中間層のユニットには実装の簡略化のために, バイアスをユニットの 1 つに見立てた特別なユニットを追加している [2]. ニューラルネットワークは式 1 および式 2 で計算されるが, バイアス b を前の層に属する 0 番目のユニットとみなすことにより, 式 1 を 式 3 のように簡略化して計算できる. なお, u_j はある層のユニット j への入力, w_{ji} はユニット j とその前の層のユニット i との結合の重み, z_i は入力層では入力値と同一であり, それ以外の層では前の層のユニット i からの出力である. f は活性化関数, z_j はユニット j の出力である.

$$u_j = \sum_{i=1}^{I} w_{ji} z_i + b_j \tag{1}$$

$$z_j = f(u_j) \tag{2}$$

$$u_j = \sum_{i=0}^{I} w_{ji} z_i \tag{3}$$

中間層における活性化関数には従来からあるロジスティクス関数や双曲線正接関数に代わって近年良く用いられる正規化線形関数を使用した. 出力層における活性化関数にはソフトマックス関数を使用したため, 誤差関数 E は式 4 のようになる. ここで W は重みとバイアスを合わせたパラメータの行列, d_k は出力ユニット k の目標値, y_k はネットワークの出力, X は入力値ベクトルである. 学習は式 4 を最小化するように確率的勾配降下法と誤差逆伝播法を用いて行った [2].

$$E(W) = -\sum_{k=1}^{K} d_k \log y_k(X; W) \tag{4}$$

2.2. 学習及び交差検定用 HTML ページ

学習及び交差検定用として, JavaScript で実装した多層ニューラルネットワークを含む HTML ページを作成した. この HTML ページには“メールアドレス”, “ID”, “パスワード”, “郵便番号”, “生年月日”の入力項目を設けた. 入力項目は後述の市販ツールが対象としているID 登録ページの入力項目を前提とした.

2.3. ツールと人における入力の違い

図 3 は 2.2. の HTML ページを利用して, ツールと人との入力時間を記録し, プロットしたものである. 中央より左がツール, 右が人による入力である. 図から分かる通り, 両者には明らかな違いが見られる.

2.4. 学習データと学習

2.2. の HTML ページを利用し, スクリプトによる自動入力 100, 手動入力 100 のデータを作成して学習した.

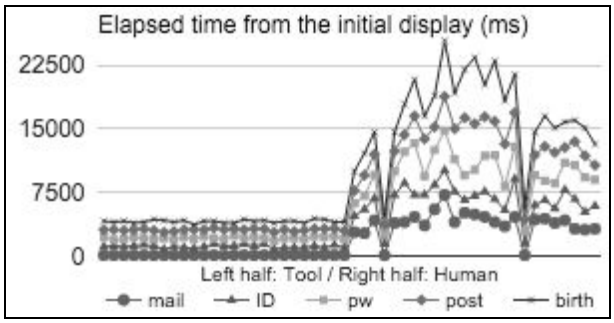

図 3: 初期表示から入力完了までの時間 (ミリ秒)

2.5. 交差検定

図 4 は 2.2. の HTML ページを利用して取得した誤差関数の値, 及び F 値を学習回数毎にプロットしたものである. 学習回数が 21 回に達した時点で誤差関数の値は 0.03, F 値は 0.97 となった.

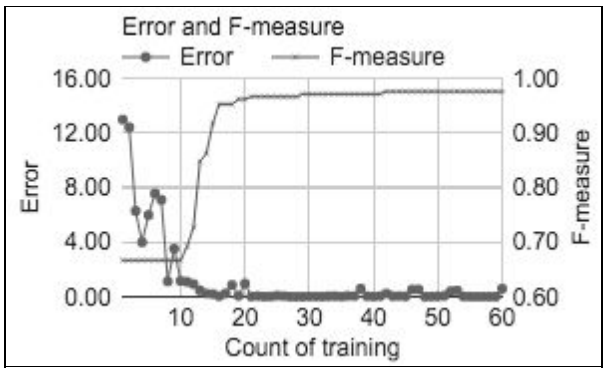

図 4: 誤差関数の値と F 値

3. 市販ツールを対象とした評価

市販されているツールを使用し, 提案手法の有効性を検証した. 表 1 は人と市販ツールによる入力を 2.4. の学習で作成したモデルを用いて判定した結果である.

表 1: 市販ツールを使用した検証結果

入力方式	試行回数	正解回数	精度
人 (手動)	10	10	100.0%
市販ツール	10	10	100.0%

4. まとめ

多層ニューラルネットワークが自動 ID 取得ツールの検知に有効であることが確認された.

参考文献

[1] 原正憲, 長谷巧, 山本匠, 山田明, & 西垣正勝. (2009). スパムブログとアフィリエイトの関連性に関する一考察. 情報処理学会論文誌, 50(12), 3206-3210.
[2] 岡谷貴之. (2015). 機械学習プロフェッショナルシリーズ 深層学習.