

ОБНАРУЖЕНИЕ БОТОВ В СОЦИАЛЬНЫХ СЕТЯХ С ПОМОЩЬЮ МНОГОСЛОЙНОГО ПЕРЦЕПТРОНА

М.Г. Хачатрян

5019973@mail.ru

SPIN-код: 3633-3934

П.И. Чепик

p.chepik@bmstu.net

SPIN-код: 3691-4451

МГТУ им. Н.Э. Баумана, Москва, Российская Федерация

Аннотация

Рассмотрена возможность использования многослойного перцептрона для обнаружения ботов в социальных сетях. Цель исследования — дать количественную оценку качества обнаружения ботов многослойным перцептроном. С помощью алгоритма кросс-валидации определены оптимальные гиперпараметры для многослойного перцептрона и значение качества классификации при найденных гиперпараметрах. Обучение и тестирование многослойного перцептрона выполнено на основе выборки из нескольких тысяч аккаунтов Twitter, состоящей как из настоящих пользователей, так и из ботов двух различных видов. В результате проведенного тестирования на этих данных получено значение метрики $F_1 = 0,958$.

Ключевые слова

Социальные сети, Twitter, бот, аккаунт, нейронные сети, кросс-валидация, метрика, классификация

Поступила в редакцию 04.04.2019

© МГТУ им. Н.Э. Баумана, 2019

Введение. За последнее десятилетие социальные сети, такие как Twitter и Facebook, хорошо зарекомендовали себя в качестве инструмента для коммуникации в реальном времени. Социальные сети затрагивают практически все слои населения и эффективно структурируют пользователей по их интересам, политическим, религиозным и иным взглядам [1]. Однако несмотря на все удобства, которые предоставляют социальные сети, существуют проблемы, связанные с их неправомерным использованием, включая политическую пропаганду, экстремистские высказывания, распространение дезинформации и т. д. [2]. Одним из способов проведения различных манипуляций в социальных сетях являются так называемые социальные боты (в дальнейшем — боты) [3]. В данной работе представлена методика обнаружения ботов в социальной сети Twitter с помощью многослойного перцептрона, а если точнее, методика определения класса аккаунта Twitter: бот или пользователь.

Многослойный перцептрон. Перцептрон (рис. 1) — элементарный нейрон, представляющий собой линейный сумматор, каждый из входных сигналов которого умножается на некоторый весовой множитель, а выходной суммарный сигнал является ненулевым, если сумма превышает некоторое пороговое значение (впервые описан в 1958 г. в работе [4]).

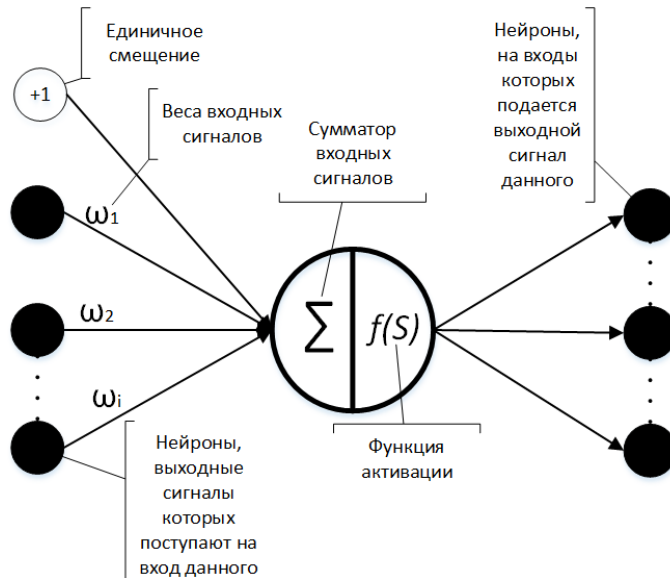


Рис. 1. Схема перцептрона

Многослойный перцептрон [5] представляет собой классический пример однонаправленной нейронной сети, состоящей из множества перцептронов. Архитектура многослойного перцептрона представлена на рис. 2.

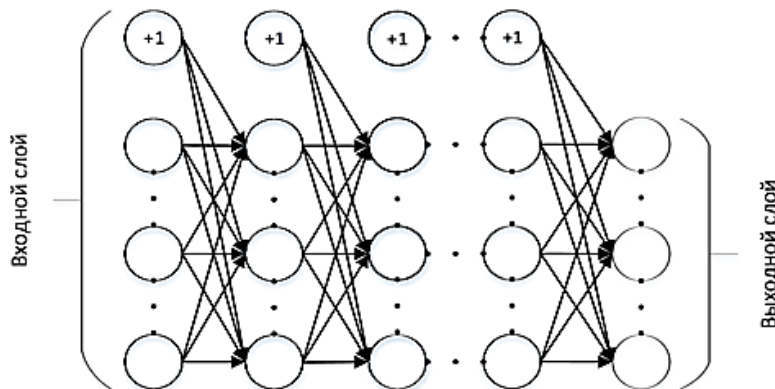


Рис. 2. Архитектура многослойного перцептрона

Размерность входного слоя определяется некоторым вектором, который содержит признаки аккаунта (такие как число друзей, подписчиков и т. д.). Размерность выходного слоя определяет число классов, однако для бинарной классификации может использоваться только один нейрон. Число скрытых слоев и количество нейронов в них определяется эмпирическим образом и является результатом многократного экспериментирования с сетью. В качестве функции активации на входном и внутренних слоях была выбрана функция ReLU [6], а для выходного слоя — сигмоидальная функция [7]. Для настройки весов нейронов (обучения сети) используется алгоритм Adam [8].

Используемые данные. Для обучения и тестирования нейронной сети были использованы данные представленные в статье [9], которые разделены на различные группы ботов и пользователей. Каждая группа — таблица, строки которой представляют собой различные аккаунты, а столбцы — признаки аккаунта (например, число групп, к которым состоит аккаунт), по которым определяется, является аккаунт ботом или нет. Описание используемых из данной базы групп представлено в табл. 1.

Таблица 1

Описание используемых групп

| Группа | Описание | Число аккаунтов |
|---|---|-----------------|
| genuine accounts | Аккаунты обычных пользователей | 3474 |
| social spambots #1 | Ретвитеры* некоего итальянского политического деятеля | 991 |
| social spambots #2 | Спамеры* платных приложений для мобильных устройств | 3457 |
| social spambots #3 | Спамеры продуктов на продажу в Amazon.com | 464 |
| Примечания: *боты, которые совершают размещения записи на своей странице для определенного аккаунта; **боты, которые распространяют спам. | | |

Методика тестирования сети. Оценка качества классификации производится методом кросс-валидации по k блокам с выбором оптимальных гиперпараметров (некоторые параметры, которые не изменяются в процессе обучения сети), описание которого представлено в статье [10]. В качестве метрики для оценки качества классификации была выбрана F_1 -мера:

$$F_1 = 2 \frac{Pre \cdot Rec}{Pre + Rec},$$

где $Pre = \frac{TP}{TP + FP}$ — точность; $Rec = \frac{TP}{TP + FN}$ — полнота; TP — число ботов, идентифицируемых как боты; TN — число пользователей, идентифицируемых как пользователи; FP — число пользователей, идентифицируемых как боты; FN — число ботов, идентифицируемых как пользователи.

Обучение и тестирование сети. Реализация и тестирование сети было произведено с помощью языка программирования Python и его библиотек. При тестировании методом кросс-валидации по 10 блокам данные были разделены на тренировочные и тестовые (90 % тренировочных, 10 % тестовых). На тренировочных данных методом кросс-валидации подбираются наилучшие гиперпараметры, а на тестовых данных дается оценка качества классификации. Зависимость F_1 -метрики от количества скрытых слоев N_l и количества нейронов N_n представлена в табл. 2.

Зависимость F_1 -метрики от количества скрытых слоев N_l и количества нейронов N_n при использовании метода кросс-валидации

| N_l | N_n | | | | | | |
|-------|-------|-------|-------|-------|-------|-------|-------|
| | 30 | 40 | 50 | 60 | 70 | 80 | 90 |
| 1 | 0,953 | 0,963 | 0,962 | 0,963 | 0,959 | 0,968 | 0,970 |
| 2 | 0,972 | 0,970 | 0,973 | 0,967 | 0,958 | 0,980 | 0,973 |
| 3 | 0,968 | 0,962 | 0,983 | 0,981 | 0,977 | 0,976 | 0,963 |

Видно, что наилучшие результаты достигаются при числе скрытых слоев $N_l = 3$ и числе нейронов $N_n = 50$. После обучения алгоритма на тренировочных данных с полученными оптимальными гиперпараметрами выполнено тестирование на тестовых данных и получено значение метрики $F_1 = 0,958$.

Выводы. В настоящей статье была описана методика обнаружения ботов в социальной сети Twitter с помощью многослойного перцептрона. Для оценки качества обнаружения ботов были использованы данные, представленные в статье [9]. Подбор оптимальных гиперпараметров многослойного перцептрона выполнен методом кросс-валидации по 10 блокам. В результате проведенной оценки точности обнаружения ботов при использовании многослойным перцептроном оптимальных гиперпараметров получено значение метрики $F_1 = 0,958$.

Литература

- [1] Алымов А.С., Баранюк В.В., Смирнова О.С. Детектирование бот-программ, имитирующих поведение людей в социальной сети «ВКонтакте». *International Journal of Open Information Technologies*, 2016, т. 4, № 8, с. 55–60.
- [2] Gayo-Avello D. Social media won't free us. *IEEE Internet Comput.*, 2017, vol. 21, no. 4, pp. 98–101. DOI: 10.1109/MIC.2017.2911439 URL: <https://ieeexplore.ieee.org/document/7994552>
- [3] Лыфенко Н.Д. Виртуальные пользователи в социальных сетях: мифы и реальность. *Вопросы кибербезопасности*, 2014, № 5(8), с. 17–20.
- [4] Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychol. Rev.*, 1958, vol. 65, no. 6, pp. 386–408. DOI: 10.1037/h0042519 URL: <https://psycnet.apa.org/doiLanding?doi=10.1037%2Fh0042519>
- [5] Рычагов М.Н. Нейронные сети: многослойный перцептрон и сети Хопфилда. *Exponenta Pro. Математика в приложениях*, 2003, № 1, с. 29–37.
- [6] Xu B., Wang N., Chen T., et al. Empirical evaluation of rectified activations in convolution network. URL: <https://arxiv.org/abs/1505.00853> (дата обращения: 15.01.2019).
- [7] Yonaba H., Ancil F., Fortin V. Comparing sigmoid transfer functions for neural network multistep ahead streamflow forecasting. *J. Hydrol. Eng.*, 2010, vol. 15, no. 4. DOI: 10.1061/(ASCE)HE.1943-5584.0000188 URL: <https://ascelibrary.org/doi/10.1061/%28ASCE%29HE.1943-5584.0000188>
- [8] Kingma D., Ba J.L. Adam: a method for stochastic optimization. *Proc. ICLR*, 2015. URL: <https://arxiv.org/abs/1412.6980> (дата обращения: 15.01.2019).

- [9] Cresci S., Pietro R.D., Petrocchi M., et al. The paradigm-shift of social spambots: evidence, theories, and tools for the arms race. *Proc. 26th Int. Conf. World Wide Web Companion*, 2017, pp. 963–972. DOI: 10.1145/3041021.3055135
URL: <https://dl.acm.org/citation.cfm?doid=3041021.3055135>
- [10] Raschka S. Model evaluation, model selection, and algorithm selection in machine learning. URL: <https://sebastianraschka.com/pdf/manuscripts/model-eval.pdf> (дата обращения: 15.01.2019).

Хачатрян Микаэл Гагикович — студент кафедры «Информационная безопасность», МГТУ им. Н.Э. Баумана, Москва, Российская Федерация.

Чепик Полина Игоревна — студентка кафедры «Информационная безопасность», МГТУ им. Н.Э. Баумана, Москва, Российская Федерация.

Научный руководитель — Ключарев Петр Георгиевич, кандидат технических наук, доцент кафедры «Информационная безопасность», МГТУ им. Н.Э. Баумана, Москва, Российская Федерация.

DETECTION OF BOTS IN SOCIAL NETWORKS BY MEANS OF MULTILAYERED PERCEPTRON

M.G. Khachatryan

5019973@mail.ru

SPIN-код: 3633-3934

P.I. Chepik

p.chepik@bmstu.net

SPIN-код: 3691-4451

Bauman Moscow State Technical University, Moscow, Russian Federation

Abstract

The possibility of using multilayer perceptron for the detection of bots in social networks is considered. The purpose of the study is to quantify the quality of detection of bots by a multilayer perceptron. Using the cross-validation algorithm, we determined the optimal hyperparameters for the multilayer perceptron and the value of the quality of classification for the found hyperparameters. The training and testing of the multilayer perceptron was carried out on the basis of a sample of several thousand Twitter accounts, consisting of both real users and bots of two different types. As a result of testing on this data the metric value is obtained $F_1 = 0,958$.

Keywords

Social networks, Twitter, bot, account, neural networks, cross-validation, metric, classification

Received 04.04.2019

© Bauman Moscow State Technical University, 2019

References

- [1] Alymov A.S., Baranyuk V.V., Smirnova O.S. Detection of bot programs that mimic the behavior of people in the social network "Vkontakte". *International Journal of Open Information Technologies*, 2016, vol. 4, no. 8, pp. 55–60 (in Russ.).
- [2] Gayo-Avello D. Social media won't free us. *IEEE Internet Comput.*, 2017, vol. 21, no. 4, pp. 98–101. DOI: 10.1109/MIC.2017.2911439 URL: <https://ieeexplore.ieee.org/document/7994552>
- [3] Lyfenko N.D. Virtual users in social networks: myths and realities. *Voprosy kiberbezopasnosti* [Voprosy kiberbezopasnosti], 2014, no. 5(8), pp. 17–20 (in Russ.).
- [4] Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychol. Rev.*, 1958, vol. 65, no. 6, pp. 386–408. DOI: 10.1037/h0042519 URL: <https://psycnet.apa.org/doiLanding?doi=10.1037%2Fh0042519>
- [5] Rychagov M.N. Neuron networks: multilayer perceptron and Hopfield nets. *Exponenta Pro. Matematika v prilozheniyakh*, 2003, no. 1, pp. 29–37 (in Russ.).
- [6] Xu B., Wang N., Chen T., et al. Empirical evaluation of rectified activations in convolution network. URL: <https://arxiv.org/abs/1505.00853> (accessed: 15.01.2019).
- [7] Yonaba H., Anctil F., Fortin V. Comparing sigmoid transfer functions for neural network multistep ahead streamflow forecasting. *J. Hydrol. Eng.*, 2010, vol. 15, no. 4. DOI: 10.1061/(ASCE)HE.1943-5584.0000188 URL: <https://ascelibrary.org/doi/10.1061/%28ASCE%29HE.1943-5584.0000188>
- [8] Kingma D., Ba J.L. Adam: a method for stochastic optimization. *Proc. ICLR*, 2015. URL: <https://arxiv.org/abs/1412.6980> (accessed: 15.01.2019).

- [9] Cresci S., Pietro R.D., Petrocchi M., et al. The paradigm-shift of social spambots: evidence, theories, and tools for the arms race. *Proc. 26th Int. Conf. World Wide Web Companion*, 2017, pp. 963–972. DOI: 10.1145/3041021.3055135
URL: <https://dl.acm.org/citation.cfm?doid=3041021.3055135>
- [10] Raschka S. Model evaluation, model selection, and algorithm selection in machine learning. URL: <https://sebastianraschka.com/pdf/manuscripts/model-eval.pdf> (accessed: 15.01.2019).

Khachatryan M.G. — Student, Department of Information Security, Bauman Moscow State Technical University, Moscow, Russian Federation.

Chepik P.I. — Student, Department of Information Security, Bauman Moscow State Technical University, Moscow, Russian Federation.

Scientific advisor — Klyucharev P.G., Cand. Sc. (Eng.), Assoc. Professor, Department of Information Security, Bauman Moscow State Technical University, Moscow, Russian Federation.