

Плекан А.І.

Львівський національний університет імені Івана Франка

Зигін С.Є.

Харківський національний університет радіоелектроніки

ВИЯВЛЕННЯ ПРОПАГАНДИ У НОВИНАХ ІЗ ВИКОРИСТАННЯМ АРХІТЕКТУРИ ДОВГОЇ КОРОТКОЧАСНОЇ ПАМ'ЯТІ

Стаття присвячена виявленню та дослідженню існування і функціонування пропаганди в умовах інтенсифікації інформації. При об'ємному потоці відомостей із різних джерел складно відокремити правдиві відомості від хибних, новини та природну інформацію від пропагандистського контенту. Більш того, більшість сучасного суспільства не замислюється над цим, для неї не характерне застосування аналітичних здібностей при сприйнятті інформації та критичне її осмислення. Отже, виникає необхідність у застосуванні комп'ютерних технологій та різних типів нейронних мереж, яке стало можливим завдяки обчислювальній потужності, що надається графічними процесорами та іншими паралельними обчислювальними платформами, більш гнучким та масштабованим програмним середовищам і великим обсягам доступних даних, що з'явилися.

В статті описано методи пропаганди, які використовуються в новинах, щоб люди могли ефективно зрозуміти їхню тему в нашому повсякденному житті. У цій статті були зібрані дані для виявлення пропаганди у новинах із використанням архітектури довгої короткочасної пам'яті (LSTM). Виявлення пропаганди розпочинається в процесі виявлення фейкових новин. З розвитком штучного інтелекту та машинного навчання, виявлення пропаганди в новинах привернуло увагу дослідників, що сприяло тому, що дана проблема стала новим напрямом. Нами було запропоновано використовувати архітектуру довгої короткочасної пам'яті (LSTM), щоб фіксувати семантичні особливості людської мови для виявлення використання пропаганди в новинних статтях. Описано алгоритм і принцип роботи даної нейронної мережі вирішено завдання класифікації та обробки даних. На основі проведеного тестування встановлено, що архітектура LSTM дозволяє ефективно викривати пропаганду в Інтернет новинах шляхом аналізу різних даних, від стилю написання та рівнів читабельності до наявності певних ключових слів.

Ключові слова: LSTM, нейронні мережі, дані, мережі, машинне навчання, засоби масової інформації.

Постановка проблеми. Пропаганда – це риторичний прийом, призначений для обслуговування певної теми, який часто цілеспрямовано використовується в новинних статтях, щоб досягти наміченої мети завдяки своєму специфічному психологічному ефекту. Існує кілька підходів до розуміння пропаганди [1]. Більшість сучасного суспільства продовжує черпати інформацію із традиційних засобів. Зокрема йдеться про телебачення, яке є найзручнішим джерелом отримання відомостей: їх не потрібно шукати, читати, аналізувати. Цю роботу вже виконали провідні телепередачі, інформація підноситься телеглядачам у готовому вигляді: знайдена, відібрана та проаналізована. Телеглядачі, як правило, беззастережно вірять у її правдивість, не піддають критичному аналізу і не намагаються знайти альтернативні погляди.

Пропаганда один із сучасних механізмів маніпулювання. У [2] наведено 5 міфів, за наявності

яких маніпуляція громадським свідомістю стає можливою:

1. Міф про нейтралітет, тобто маніпуляція повинна залишатися непоміченою, люди повинні вірити у те, що інформація природна, а державні та політичні інститути нейтральні і не відстоюють власні інтереси. Найбільш «нейтральними» у цій хибній реальності повинні здаватися верховні державні інститути та ЗМІ.

2. Міф про плюралізм ЗМІ, тобто створення ілюзії про те, що вибір джерела інформації існує. Інакше кажучи, створюється міф у тому, що велика різноманітність інформаційних каналів транслюють різний за змістом контент. Насправді джерело контенту і власник численних інформаційних каналів один й той самий, тому сенс переданої інформації однаковий.

3. Міф про відсутність соціальних конфліктів. Маніпулятори створюють ілюзію повної соціаль-

ної гармонії в країні, представляючи насильство та конфлікти як окремі випадки.

4. Міф про індивідуалізм та особистий вибір. Приватництво і підприємництво породили міф про індивідуальність людини та її незалежність від суспільства. Однак насправді, за словами Г. Шіллера, людина не відокремлена від суспільства.

5. Міф про постійну природу людини. Іншими словами, людині нав'язується концепція, наприклад, про агресивну та невинуватну природу людини.

Таким чином, управління суспільною свідомістю за допомогою пропаганди як механізму маніпуляції може здійснюватися різними методами. Також свідомістю має простіше керувати, якщо їх переконати, що деякі концепції та розумові шаблони є істинними та природними. Особливо небезпечною пропаганда стає під час військових конфліктів коли вона використовується ворогом для деморалізації суспільства, створення негативного сприйняття владних рішень, зриву мобілізації тощо. Тому актуальною є проблема виявлення пропаганди для того щоб ефективно використовувати це в нашому повсякденному житті.

Аналіз останніх досліджень і публікацій. На думку Г. Лассуелла [3], пропаганда – це управління колективними установками у вигляді маніпулювання значними символами. Основна функція пропаганди – усувати соціальну дезорганізацію шляхом переконання та аргументів, а не насильства та тиранії. Демократія, на його думку, проголосила диктатуру демагогії, а диктатором – пропаганду. У [4] автори дотримуються визначення пропаганди, даного американським психологом У. Байдлом: під впливом пропаганди кожен індивід поводить себе так, начебто його поведінка впливає з його рішень. У [5] описано історію походження пропаганди та ставлення до неї, що сформувалося в Європі.

Визначення пропаганди як сучасний вплив правлячої меншини на маси, за допомогою якого правляча еліта формує думку мас і спрямовує її сили наведено в [6]. При цьому зазначено, що масова освіта не виконала цілі, які стояли перед нею, і не зробили людей критично мислячими. За їх допомогою маси набули лише набір штампів, кліше та стереотипів, які полегшують проведення пропагандистських кампаній. На думку автора, пропаганда є комбінованим (поєднує інформаційне та фізичне) видом впливу на людину або маси, метою якої є збіль-

шення впливу та влади. У [7] зазначено, що пропаганда є найважливішою технологією регулювання суспільних відносин, що здійснюється за допомогою формування суспільних установок та їх бажаних поєднань. На думку автора, для проведення вдалої пропагандистської кампанії необхідно ґрунтуватися на наукових здобутках педагогіки, психології, соціології.

Питанню виявлення пропаганди присвячено ряд наукових публікацій. Одним з методів, що використовується дослідниками в цьому напрямку є виявлення рівнів речення. У [8] запропоновано модель для визначення рівня речення, яка могла б зрозуміти семантичні особливості покращення мови шляхом побудови контекстно-залежного введення пар (пара речення-заголовок і пара речення-контекст). Натомість у [9], як інструменти для виявлення пропаганди, були досліджені різні архітектури нейронних мереж (CNN, LSTM-CRF і BERT). Загалом, зазначено, в [10-11], найпопулярнішими методами серед методів глибокого навчання є нейронні мережі, серед яких можна виділити рекурентні нейронні мережі (RNN), довгу короткочасну пам'ять (LSTM), двонаправлену довгу короткочасну пам'ять (Bi-LSTM), згорткові нейронні мережі (CNN) та інші. Крім того, багато досліджень використовують методи машинного навчання для виявлення джерел дезінформації, такі як метод опорних векторів (SVM), наївний класифікатор Байєса (NBC), випадковий ліс (RF), дерева рішень (DT), логістична регресія (LR) і Баєсове моделювання (BM). На основі порівняння було виявлено, що методи на основі машинного навчання більш ефективні та використовуються багатьма дослідниками.

Постановка завдання. Метою статті є дослідження використання архітектури довгої короткочасної пам'яті для виявлення пропаганди у новинах.

Виклад основного матеріалу. Як зазначено вище, одним з інструментів для виявлення пропаганди можуть бути нейронні мережі. Так, наприклад, у [10] розглянуто застосування нейронних мереж останнього покоління для виявлення пропаганди, а в [11] проведено розрахунок на тришарових нейронних мережах. Результати розрахунку залежності класифікації від епохи навчання при вибраному оптимальному параметрі швидкості навчання для кожного методу представлені в табл. 1. Розрахунки авторами [11] проводились на тришарових нейронних мережах з розміром шарів 63, 31 та 16 відповідно. Нейронна мережа, навчена методом адаптивного моменту (Adam), має більш

високу точність класифікації і досягає 93% на сотій епосі. Нейронні мережі, навчені іншими методами, що досягають точності 85–90% [15].

Таблиця 1

Епохи навчання [11]

Метод навчання	Оптимальна швидкість навчання
GD	0.1
qProp	0.05
NAG	0.5
AdaGrad	0.1
AdaDelta	0.05
Adam	0.1

LSTM або довга короткострокова пам'ять – це особливий різновид рекурентної нейронної мережі (RNN – Recurrent Neural Network), здатний до навчання довгостроковим залежностям. Вперше вони були представлені 1997 року вченими Зеппом Хохрайтером та Юргеном Шмідхубером [16]. Архітектура мережі розгортається для відображення всієї мережі у вигляді повної послідовності. Мережі довгої короткострокової пам'яті розроблені для того, щоб уникати проблем довгострокової залежності. Запам'ятовування інформації та її зберігання в пам'яті – є стандартною поведінкою для LSTM.

Стан комірки – горизонтальна лінія, що проходить через верх діаграми, є ключем до LSTM. Її стан схожий на конвеєрну стрічку з лінійними взаємодіями. Для потоку інформації не важко переміщатися по комірці без змін. Однією з особливостей є здатність додавати інформацію про стан комірки, суворо регульованими структурами, званіми вентилями. Вентилі мають можливість пропустити інформацію через себе за бажанням.

Мережі з короткочасною пам'яттю являють собою видозмінену версію періодичних нейронних мереж, які у свою чергу спрощують процес запам'ятовування минулих даних у пам'яті. Також LSTM добре підходить для прогнозування тимчасових рядів, де враховуються тимчасові затримки невідомої тривалості. Навчання моделі відбувається за допомогою зворотного поширення помилки. На основі LSTM нами досліджено рекурсивну формулу мережі RNN:

$$h_t = \tanh(W_h h_{t-1} + W_x x_t) \quad (1)$$

$$y_t = W_y h_t \quad (2)$$

де x_t – вхідний вектор, h_t – прихований шар, y_t – вектор виведення експерименту, W_h – зважена матриця.

Рекурентні нейронні мережі застосовуються до довгої короткострокової пам'яті для створення процесу обчислень, отриманих вхідних даних та створення вихідних даних [12]. У ході цього процесу довготривала пам'ять створюється з короткострокової. Система LSTM складається з шару вхідного вентиля (input gate), вентиля забування (forget gate), та вихідного вентиля (output gate). Архітектура LSTM показана на рис. 1 [13]. LSTM обчислює прихований стан наступним способом:

$$i_t = \sigma(W_f[h_{t-1}, x_t] + b_i) \quad (3)$$

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (4)$$

$$\sigma(x) = \frac{1}{1+e^{-x}} \quad (5)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (6)$$

$$c_{\sim t} = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (7)$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot c_{\sim t} \quad (8)$$

$$h_t = o_t \cdot \tanh(c_t) \quad (9)$$

де σ – логістична сигмоїдальна функція, i , f , o – вхідні вентиля. h – прихований вектор одного розміру в кожному шарі, W – матриця для перетворення інформації та $[h_{t-1}, x_t]$ – у кожному вентилі. У рівнянні $c_{\sim t}$ – прихований елемент, який є вхідним шаром, c_t – внутрішня пам'ять, що обчислюється у пристрої, і h_t – вихід прихованого стану.

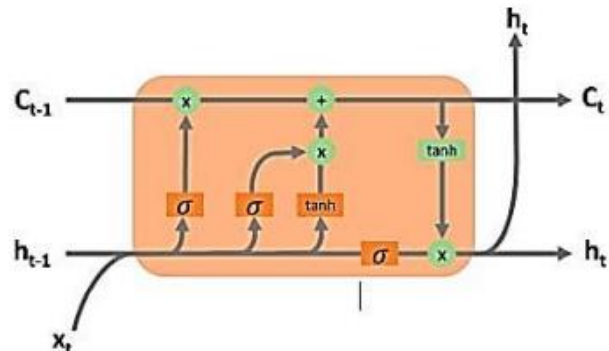


Рис. 1. Архітектура LSTM [13]

Вентиль відповідає за видалення інформації з комірки. Він отримує вихід прихованого стану з попереднього тимчасового кроку (h_{t-1}) та вхід поточного тимчасового кроку. Вхідні дані множаться на вагові матриці. Вхідний вентиль передає інформацію про стан комірки. Сигмоподібна функція застосовується як фільтр для h_{t-1} і x_t .

Вихідний вентиль вирішує яку інформацію виводити з комірки. Цей етап виконується за три кроки. Для масштабування стану комірки від 0

до 1 будують вектор, а відфільтровані значення множаться на початковий вектор, для отримання вихідної інформації. Вентилі LSTM показані рис. 2 [14].

Для того, щоб використовувати LSTM модель для виявлення пропаганди у новинах необхідно:

- Очистити та впорядкувати дані. Дані поділяються на два стовпці. У першому це кількість випадків пропаганди за індексом якості визначення та у другому – в який день;
- Розбити дані на частини: навчання, перевірка, прогнозування;
- Перенести проблему часових рядів у контрольоване завдання.

Основним напрямом у вивченні нейронних мереж є дослідження залежності між типом нейронної мережі, її архітектурою та результатами її роботи. Для покращення продуктивності моделі прийнято використовувати середньоквадратичну помилку та кореляцію. Як вхідні дані для тренування моделей використовувалися дані з оптимальними параметрами [епоха, швидкість навчання (lr), довжина речення ($sentlen$)] для моделей двох завдань SI ($span\ identification$ – ідентифікація фрагменту) та TC ($technique\ classification$ – класифікація техніки.). Порівняння різних моделей наведено в таблиці 2, де базовий рівень визначався за допомогою LSTM

Таблиця 2

Порівняння моделей
для виявлення пропаганди

Модель	F1*
Базовий рівень (Baseline)	0.262326
BERT	0.539535
Наша система ТК	0.575729

*F1-бал визначає пропаганду

Для моделі прогнозу з TC використовувався алгоритм векторизації текстових даних LSTM (для завдання SI), результати наведені на рис. 3. Виходячи з наведених вище оптимальних пара-

метрів, системи SI і TC отримали оцінку F1 0,441732 і 0,575729, і обидва процеси навчання зайняли приблизно 2,5 години з використанням 4 відеокарт Nvidia GTX 1080Ti (тобто близько 10 годин GPU). В [10], обчислення відбувалося на відеокарті NVIDIA A100 з показником 0,320 за 3 години роботи. Наведені дані F1 трохи перевищують середньостатистичні згідно методики описаної в [11]. Результати показують, що моделі загалом перевершують інші моделі для завдання ідентифікації діапазону пропаганди.

Проведено ручне тестування моделі для ідентифікації фейкових новин, використовуючи новини з набору даних, які не використані під час навчання. Усі новини використані для ручної перевірки були розбиті на тренувальну та тестову множину, в результаті перевірки більшість новин ідентифіковано правильно з точністю 90%. Перспективи подальших досліджень полягають у розширенні набору даних із різних джерел для покращення процесу опрацювання тексту.

Висновки. В даній роботі розглянуто проблему використання пропаганди в новинах і повсякденному житті людей. Використано базуючись на моделі LSTM дві конкретні системи ідентифікації діапазону та класифікації техніки пропаганди в новинних статтях. Для завдання SI ми заново визначили його як завдання тегування послідовності з трьома маркерами за допомогою системи SI і прийнятого методу конкатенації ознак на рівні речення. Досліджено на основі LSTM ефективні і точні системи виявлення пропаганди в новинах. Це сприятиме реалізації інтегрованих систем для автоматичного виявлення фейкових новин у реальному часі для розпізнавання маніпулятивних і тенденційних новин. Ці перспективи даватимуть змогу не тільки підвищити точність та ефективність моделі, але й зробити її більш універсальною та придатною для використання в реальних умовах ведення інформаційної війни з росією в інтернет-просторі.

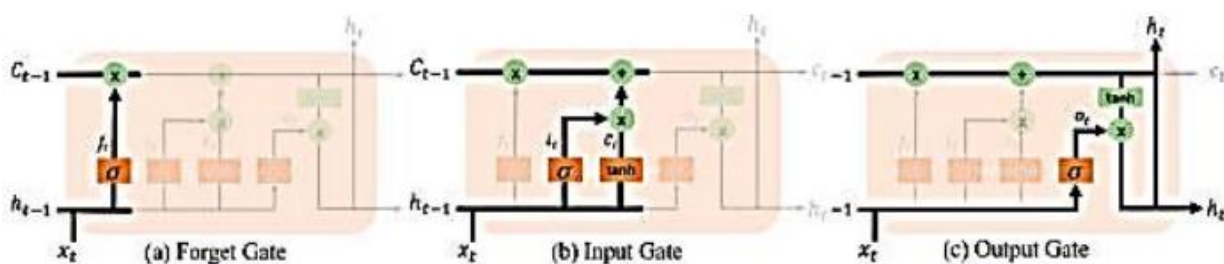


Рис. 2. Вентилі LSTM

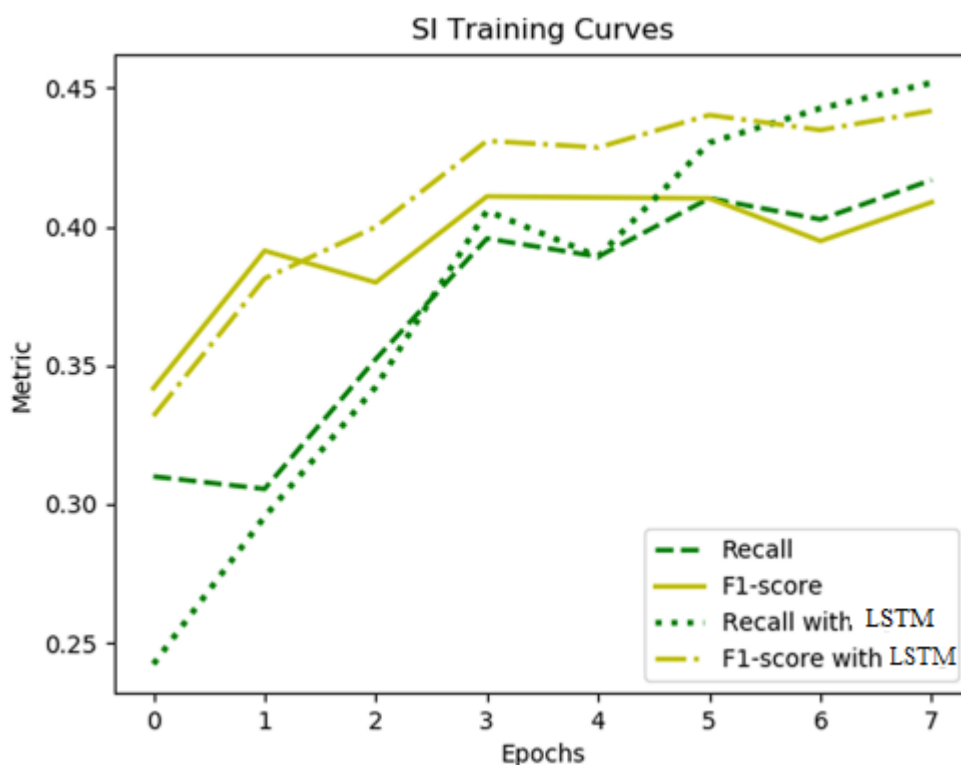


Рис. 3. Порівняння SI навчального процесу визначення пропаганди F1 з і без LSTM

Список літератури:

1. Haykin S. Neural Networks. Second Edition, Addison Wesley Longman, 2009. 485 p.
2. Barron-Cedeno A, Da San Martino G, Jaradat I, Nakov P. Propgy: a system to unmask propaganda in online news. In: Proceedings of the AAAI conference on artificial intelligence, 2019, 9847–9848.
3. Lasswell H.D. The Theory of Political Propaganda. The American Political Science Review, 1927, 21(3), 627–631.
4. Zhang Z, Sabuncu M. Generalized cross entropy loss for training deep neural networks with noisy labels. In: Advances in neural information processing systems, 2018, 8778–8788.
5. Yoosuf S., Yang Y. Fine-grained propaganda detection with fine-tuned bert. In: Proceedings of the second workshop on natural language processing for internet freedom: censorship, disinformation, and propaganda, 2019, 87–91.
6. Edward Bernays. Propaganda. – [Electronic resource]. – URL: <http://commondatastorage.googleapis.com/brainmod/book/EdwardBernays-Propaganda.pdf>. pp. 4-5. Date of access: 17.12.2023.
7. Jordan M. I. Serial order: a parallel distributed processing approach // Advances in Psychology. 1997, 121, 471–495.
8. Lan Z, Chen M, Goodman S, Gimpel K, Sharma P, Soricut R. Albert: a lite bert for self-supervised learning of language representations. ICLR, 2020.
9. Liu S, Lee K, Lee I. Document-level multi-topic sentiment classification of email data with bilstm and data augmentation. Knowl Based Syst:105918, 2020.
10. Mapes N, White A, Medury R, Dua S. Divisive language and propaganda detection using multi-head attention transformers with deep learning bert-based language models for binary classification. In: Proceedings of the second workshop on natural language processing for internet freedom: censorship, disinformation, and propaganda, 2019, 103–106
11. Martino DSG, Yu S, Barron-Cedeno A, Petrov R, Nakov P. Fine-grained analysis of propaganda in news articles. EMNLP/IJCNLP, 2019, 1:5635–5645
12. Pankaj G, Khushbu S, Usama Y, Thomas R, Hinrich S. Neural architectures for fine-grained propaganda detection in news. In: Proceedings of the second workshop on natural language processing for internet freedom: censorship, disinformation, and propaganda, 2019.

13. San G.D.M, Alberto B.C, Preslav N. Findings of the nlp4if-2019 shared task on fine-grained propaganda detection. In: Proceedings of the second workshop on natural language processing for internet freedom: censorship, disinformation, and propaganda, 2019.
14. Tchiehe N.D, Gauthier F. Classification of risk acceptability and risk tolerability factors in occupational health and safety. Saf Sci:138–147, 2017.
15. Wang A, Singh A, Michael J, Hill F, Levy O, Bowman R.S. Glue: A multi-task benchmark and analysis platform for natural language understanding. In: International conference on learning representations, 2018.
16. S. Hochreiter and J. Schmidhuber, “Long short-term memory,” Neural Computation, 1997.

**Plekan A.I., Zyhyn S.Ie. DETECTING PROPAGANDA IN THE NEWS
USING THE ARCHITECTURE OF LONG SHORT-TERM MEMORY**

The article is devoted to the detection and study of the existence and functioning of propaganda in the conditions of information intensification. With a voluminous flow of information from various sources, it is difficult to separate true information from false, news and natural information from propaganda content. Moreover, most of modern society does not think about this, it is not typical for it to use analytical abilities in the perception of information and its critical understanding. Therefore, there is a need to use computer technologies and various types of neural networks, which became possible due to the computing power provided by graphics processors and other parallel computing platforms, more flexible and scalable software environments and large volumes of available data that have appeared.

The article selects propaganda methods used in the news so that people can effectively understand their topic in our everyday lives. This article collected data on the detection of propaganda in the news using the long short-term memory (LSTM) architecture. The detection of propaganda methods begins with the process of detecting fake news. With the development of artificial intelligence and machine learning, the detection of propaganda in the news has attracted the attention of researchers, which has contributed to the fact that this problem has become a new direction. We proposed to use the architecture of long short-term memory (LSTM) to capture the semantic features of human language to detect the use of propaganda in news articles. The algorithm and the principle of operation of this neural network are described, the task of classification and data processing is solved. Based on the testing conducted, it was established that the LSTM architecture allows you to effectively expose propaganda in Internet news by analyzing various data, from writing style and readability levels to the presence of certain keywords.

Key words: LSTM, neural networks, data, networks, machine learning, mass media.