

Контроль космічного та повітряного простору

УДК 621.384

doi: 10.26906/SUNZ.2018.5.020

Є. О. Гришманов

Льотна академія Національного авіаційного університету, Кропивницький, Україна

ВИБІР МАТЕМАТИЧНОГО АПАРАТУ ДЛЯ ПОБУДОВИ МОДЕЛІ ПРОГНОЗУВАННЯ НЕСПРИЯТЛИВИХ АВІАЦІЙНИХ ПОДІЙ ПІД ЧАС ПОЛЬОТУ

Мета статті. Проведення дослідження та вибір найбільш ефективного математичного апарату для побудови моделі прогнозування несприятливих авіаційних подій під час польоту. **Результати.** В статті проведений аналіз відомих методів, що використовуються для вирішення задач класифікації даних з точки зору доцільності їх застосування для побудови моделі прогнозування несприятливих авіаційних подій під час польоту на основі аналізу текстових повідомлень. Розглянуто наступні методи: логістична регресія, метод опорних векторів, наївний байєсівський класифікатор, випадковий ліс (random forest). Крім того для вирішення подібного класу задач розглянуто згорткові та рекурентні нейронні мережі в яких застосовуються алгоритми глибокого навчання. **Висновки.** В результаті аналізу вказаних методів для побудови моделі прогнозування несприятливих авіаційних подій під час польоту на основі аналізу текстових повідомлень обрано математичний апарат глибоких нейронних мереж. Завдяки застосуванню в них алгоритмів глибокого навчання вони володіють найбільш високою точністю у порівнянні з традиційними підходами.

Ключові слова: прогнозування, класифікація, байєсівський класифікатор, метод опорних векторів, згорткова нейронна мережа, рекурентна нейронна мережа.

Вступ

Постановка проблеми. Незважаючи на стрімкий науково-технічний прогрес, що спостерігається в останні десятиліття у галузі авіаційної техніки, проблема забезпечення безпеки польотів не втрачає своєї актуальності [1]. Постійна модернізація та створення новітніх зразків літальних апаратів, розробка та впровадження нових методів експлуатації авіаційної техніки, вдосконалення підготовки льотної складу та осіб, що забезпечують польоти обумовлює необхідність вдосконалення методів управління процесами забезпечення безпеки польотів. Одним з шляхів підвищення ефективності роботи системи управління безпекою польотів є застосування автоматизованих систем прогнозування несприятливих авіаційних подій під час польоту.

Традиційним підходом при розробці методів прогнозування авіаційних подій є застосування статистичних методів оцінки ризиків та факторів небезпеки. Незважаючи на ряд переваг, дані методи не завжди є універсальними і потребують достатньої кількості аналітичної та статистичної інформації стосовно об'єкту, що аналізується.

Перспективним напрямком є технології, що засновані на використанні елементів штучного інтелекту, а саме нейронних мереж. Їх актуальність обумовлена тенденцією збільшення складності математичних та формальних моделей реальних систем та необхідністю урахування великої кількості факторів, що потребує великих затрат часу та засобів.

Мета статті. Проведення дослідження та вибір найбільш ефективного математичного апарату для побудови моделі прогнозування несприятливих авіаційних подій під час польоту.

Основний матеріал

У загальному вигляді задача прогнозування несприятливих авіаційних подій під час польоту полягає у побудові класифікатора, що буде прогнозувати появу визначеного класу авіаційної події. Процес прогнозування виникнення авіаційної події деякого класу в даній роботі пропонується розглядати як задачу аналізу текстових повідомлень, що формуються за результатами збору даних від різних джерел інформації в процесі управління повітряним рухом.

Побудову моделі прогнозування несприятливих авіаційних подій під час польоту на основі аналізу текстових повідомлень можливо здійснювати з використанням наступних методів: логістична регресія, метод опорних векторів, наївний байєсівський класифікатор, випадковий ліс (random forest). Проведемо аналіз вказаних методів з точки зору доцільності їх застосування для побудови моделі прогнозування несприятливих авіаційних подій під час польоту на основі аналізу текстових повідомлень.

Логістична регресія [2] – статистична модель, за допомогою якої здійснюється передбачення ймовірності виникнення деякої події в залежності від значення множини ознак. Суттєвим недоліком даного методу є те, що він працює з обмеженою кількістю вхідних даних, тобто існує загроза не врахування деяких важливих факторів, що спричинить суттєвий вплив на результат.

Метод байєсівської класифікації [3] традиційно використовується для вирішення задач класифікації тексту. Засновується на застосуванні теореми Байєса зі строгими припущеннями про незалежність елементів вектора ознак. Перевагою наївного байєсівсько-

го класифікатора є мала кількість даних для навчання, необхідних для оцінки параметрів, низькі обчислювальні витрати та простота реалізації. Серед недоліків слід відмітити відносно невисоку якість класифікації даних при вирішенні завдань.

Випадковий ліс (random forest) [2] – алгоритм, що полягає у використанні ансамблю (комітету) дерев прийняття рішень. Класифікація об'єктів проводиться шляхом голосування: кожне дерево комітету відносить об'єкт, що класифікується до одного з класів. Перемагає той клас, за який проголосувала найбільша кількість дерев. Оптимальна кількість дерев обирається таким чином, щоб мінімізувати помилку класифікатора на тестовій вибірці. Використовує усереднення для підвищення точності прогнозування і контролю надлишкової підгонки. Перевагами даного методу є достатньо висока точність, відсутність проблеми перенавчання, простота застосування. Недоліками є відсутність візуального представлення процесу прийняття рішень та складність їх інтерпретації.

Метод опорних векторів (SVM) [2] полягає у переводі вихідних векторів у простір більш високої розмірності та у пошуку розділяючої гіперплощини для розділення позитивних та негативних прикладів з максимальним зазором в цьому просторі. Перевагою даного методу є висока ефективність та якість класифікації у порівнянні з іншими методами. Серед недоліків слід зазначити необхідність великого обсягу пам'яті та витрат машинного часу на навчання, низьку швидкість навчання та зниження ефективності методу при збільшенні кількості ознакових описів.

Для вирішення подібного класу задач ефективним на даний момент є використання глибоких нейронних мереж, які завдяки застосуванню алгоритмів глибокого навчання володіють найбільш високою точністю у порівнянні з традиційними підходами.

Глибокими нейронними мережами [4] називають мережі, що мають декілька прихованих шарів, як правило від 4-х і вище. Кожний прихований шар обчислює нелінійне перетворення попереднього шару, отже глибока нейронна мережа має значно більшу репрезентативну потужність, що дозволяє представляти значно складніші функції, ніж нейронна мережа з 1–3 шарами. При цьому в процесі навчання глибокої нейронної мережі важливо використовувати нелінійну функцію активації в кожному прихованому шарі. Таким чином, однією з найголовніших переваг глибоких нейронних мереж є стисле представлення досить великої кількості функцій.

Згортова нейронна мережа (convolutional neural network (CNN)) – клас глибоких нейронних мереж, що спочатку була створена для ефективного розпізнавання цифрових зображень [5]. Класична архітектура CNN являє собою чергування згорткових шарів з нелінійними активаційними функціями та шарами об'єднання. На відміну від мережі прямого поширення, де кожен вхідний нейрон з'єднується з вихідним нейроном в наступному шарі, в згорткових мережах для отримання вихідних значень використовуються згортки над кожним вхідним шаром. В операції згортки використовується матриця ваг

невеликого розміру, яка зсувається по всьому шару, що обробляється, формуючи після кожного зсуву сигнал активації для нейрона наступного шару з аналогічною позицією. Ця матриця називається ядром згортки та використовується для різних нейронів вхідного шару.

Безпосередньо під час аналізу текстових повідомлень на вхід CNN подається матриця, кількість рядків якої залежить від розмірності словника, а ширина фільтрів дорівнює кількості стовпців цієї матриці (розмірності, що використовуються для кодування кожного слова). Висота (розмір фрагменту вхідних даних) може змінюватися та зазвичай становить близько 2-5 слів. Перші шари являють собою слова в вигляді низькорозмірних векторів. Наступний шар виконує згортки над векторними представленнями слів, використовуючи фільтри різних розмірів (як правило, захоплюють 3-5 слів одночасно). Потім проводиться пулінг над результатом згортки. До отриманого довгому вектору ознак застосовується регуляризація. На завершальному етапі проводиться класифікація результату за допомогою шару softmax [4] (рис. 1).

На даний момент CNN вважаються досить ефективними насамперед для аналізу неструктурованого тексту. У контексті вирішення задачі прогнозування несприятливих авіаційних подій під час польоту такі мережі ефективні при використанні в процесі функціонування гібридної нейромережевої моделі прогнозування несприятливих авіаційних подій під час польоту, у випадку, коли на вхід системи надходять неструктуровані дані. При цьому у якості векторного шару використовується попередньо навчений шар Embedding [5] на структурованих навчальних даних.

Рекурентні нейронні мережі [6] (recurrent neural network (RNN)) – клас глибоких нейронних мереж зі зворотними зв'язками, які використовують попередні стани мережі для обчислення поточного стану. Найпростіший приклад такої мережі наведено на рис. 2.

Традиційно даний клас нейронних мереж використовується для задач обробки послідовностей нефіксованої довжини, зокрема для розпізнавання текстів та мови.

Під час аналізу тексту в RNN на кожному часовому кроці t вихідні дані попереднього кроку разом з наступним вектором слова x_t тексту, являють собою вхідні дані для прихованого шару для створення передбачення \hat{y}_t та ознак h_t

$$h_t = \sigma\left(W^{(hh)}h_{(t-1)} + W^{(hx)}x_t\right) \quad (1)$$

$$\hat{y}_t = \text{softmax}\left(W^{(s)}h_t\right), \quad (2)$$

де $x_{t1}, \dots, x_t, \dots, x_T$ – вектори слів словника T ; h_t – вихідна ознака прихованого шару на кожному часовому кроці t ; $x_t \in R^d$ – слово, подане на вхід в момент часу t ; $W^{(hx)} \in R^{D_h \times d}$ – матриця ваг для обробки вхідного слова x_t ; $W^{(hh)} \in R^{D_h \times D_h}$ – матриця ваг для обробки

вихідних даних на попередньому часовому кроці h_{t-1} ; $h_{t-1} \in R^{D_h}$ – вихід нелінійної функції на попередньому часовому кроці t , h_0 – вектор ініціалізації

для прихованого шару в момент часу $t = 0$; σ – нелінійна функція; \hat{y}_t – вихідний розподіл ймовірності за словником на кожному часовому кроці t .

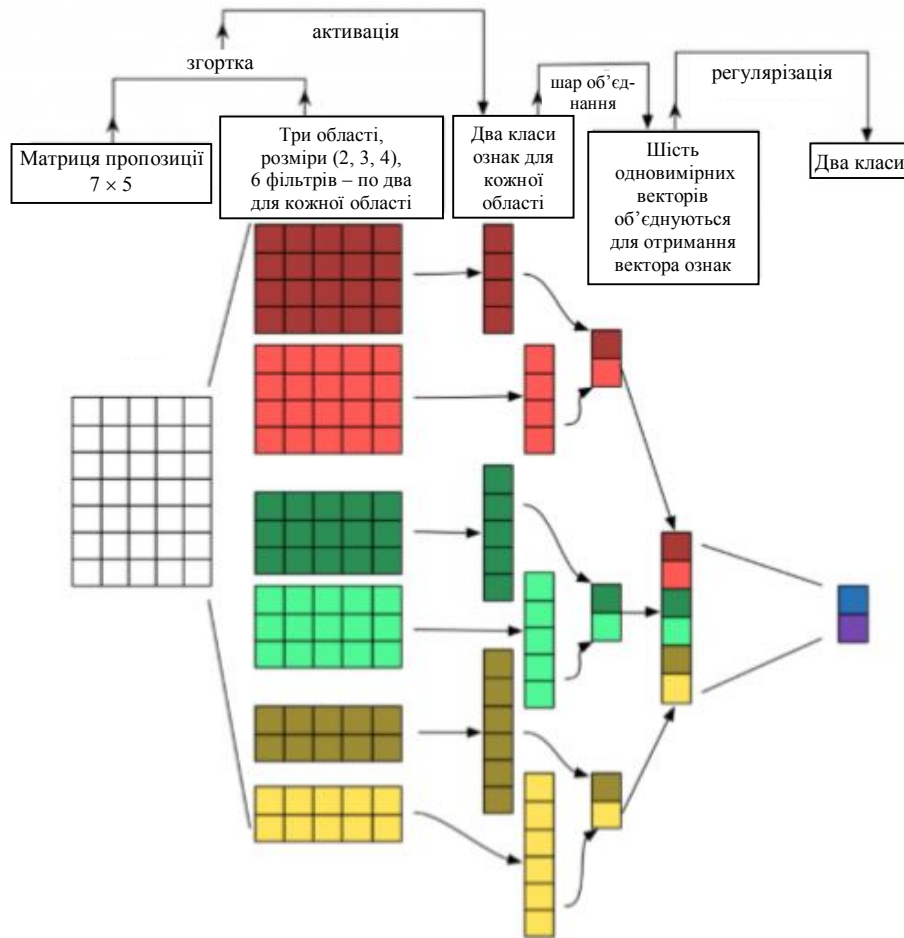


Рис. 1. Узагальнена структура топології та схема функціонування CNN для аналізу текстів

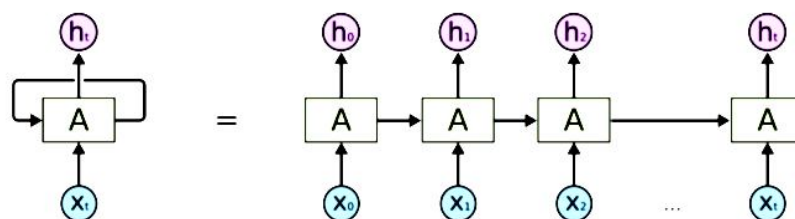


Рис. 2. Найпростіша архітектура рекурентної нейронної мережі з одним вхідним нейроном, одним прихованим нейроном, одним вихідним нейроном та її розгорнуте часове представлення

Недоліком RNN є зникнення або перевищення градієнтів в процесі навчання з використанням методу зворотного розповсюдження помилки. Використання архітектури модулів довгої короткострокової пам'яті (Long Short Term Memory (LSTM)) [7] дозволяє уникнути проблеми зникаючого градієнта в рекурентних нейронних мережах. Схема такої мережі представлена на рис. 3.

Всі RNN мають форму ланцюга блоків нейронної мережі, що повторюються. LSTM мають подібну ланцюгову структуру, однак блок який повторюється має 4 шари мережі, що впливають один на одного. RNN на базі модулів LSTM вважаються ефективними насамперед при аналізі структурованого тексту.

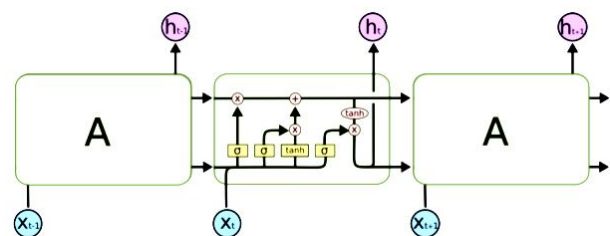


Рис. 3. Архітектура довгої короткострокової пам'яті LSTM, розгорнутої у часі

У контексті вирішення завдання прогнозування несприятливих авіаційних подій в польоті дані мережі ефективні для першого налаштування шару Embedding на структурованих навчальних даних в

процесі навчання гібридної нейромережевої моделі прогнозування несприятливих авіаційних подій під час польоту.

Аналіз вищенаведених методів показав, що найбільш доцільним для побудови моделі прогнозування несприятливих авіаційних подій під час польоту є застосування глибоких нейронних мереж.

Висновки

В роботі проведено аналіз особливостей, переваг та недоліків ряду відомих методів, що можуть

застосовуватися для вирішення задачі класифікації даних.

В результаті проведеного аналізу було встановлено, що для побудови моделі прогнозування несприятливих авіаційних подій в польоті на основі аналізу текстових повідомлень найбільш доцільним є використання глибоких нейронних мереж, а саме згорткових та рекурентних нейронних мереж.

Завдяки застосуванню в них алгоритмів глибокого навчання вони володіють найбільш високою точністю у порівнянні з традиційними підходами.

СПИСОК ЛІТЕРАТУРИ

1. Состояние безопасности полетов в гражданской авиации государств-участников соглашения о гражданской авиации и об использовании воздушного пространства в первом полугодии 2018 г. [Электронный ресурс]. – Режим доступа: <https://mak-iac.org/upload/iblock/5b0/bp-18-1.pdf>.
2. П.Флах Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных / пер. с англ. А.А.Слинкина. – М.: ДМК Пресс, 2015. – 400 с.
3. Бідюк П.І. Основні етапи побудови і приклади застосування мереж Байеса / П.І. Бідюк, Н.В. Кузнецова // Систем. дослідж. та інформ. технології. — 2007. — № 4. — С. 26 — 39.
4. dos Santos C.N., Gatti M. Deep Convolutional Neural Networks for Sentiment Analysis of Short Texts // COLING, 2014. P. 69 – 75.
5. Y. Kim. Convolutional neural networks for sentence classification. arXiv:1408.5882 [cs.CL], 2014.
6. C. Olah. Neural networks, recurrent neural networks, convolutional neural networks [Электронный ресурс]. – Режим доступа: <http://colah.github.io>.
7. Hochreiter S., Schmidhuber J. Long short-term memory // Neural computation 9, 1997. Issue 8. P. 1735 – 1780.

Рецензент: д-р техн. наук, проф. О. І. Тимочко,

Харківський національний університет Повітряних Сил імені Івана Кожедуба, Харків

Received 22.06.2018

Accepted for publication 12.09.2018

Выбор математического аппарата для построения модели прогнозирования неблагоприятных авиационных происшествий в полете

Е. А. Гришманов

Цель статьи. Проведение исследования и выбор наиболее эффективного математического аппарата для построения модели прогнозирования неблагоприятных авиационных происшествий во время полета. **Результаты.** В статье проведен анализ известных методов, используемых для решения задач классификации данных с точки зрения целесообразности их применения для построения модели прогнозирования неблагоприятных авиационных событий в полете на основе анализа текстовых сообщений. Рассмотрены следующие методы: логистическая регрессия, метод опорных векторов, наивный байесовский классификатор, случайный лес (random forest). Кроме того, для решения подобного класса задач рассмотрены сверточные и рекуррентные нейронные сети, в которых применяются алгоритмы глубокого обучения. **Выводы.** В результате анализа указанных методов для построения модели прогнозирования неблагоприятных авиационных происшествий во время полета на основе анализа текстовых сообщений выбран математический аппарат глубоких нейронных сетей. Благодаря применению в них алгоритмов глубокого обучения они обладают наиболее высокой точностью по сравнению с традиционными подходами.

Ключевые слова: прогнозирование, классификация, байесовский классификатор, метод опорных векторов, сверточная нейронная сеть, рекуррентная нейронная сеть.

Choice of mathematical instrument for model of forecasting of unfavorable aircraft accidents in the flight

E. Gryshmanov

The purpose of the article. Carrying out research and selection of the most effective mathematical device for constructing the model of forecasting of adverse aviation events during the flight. **Results.** The paper analyzes the known methods used to solve data classification problems. This is necessary to determine the feasibility of their use for building a model of forecasting unfavorable aircraft accidents in the flight, based on the analysis of text messages. The following methods are considered: logistic regression, support vector machine, naive Bayes classifier, random forest. In addition, to solve this class of problems, convolutional and recurrent neural networks using deep learning algorithms are considered. **Conclusions.** As a result of the analysis of these methods, a mathematical instrument of deep neural networks was chosen to build a model of forecasting of unfavorable aircraft accidents in the flight, based on the analysis of text messages. Due to the application of their deep learning algorithms, they have the highest accuracy compared to traditional approaches.

Keywords: forecasting, classification, Bayes classifier, support vector machine, convolutional neural network, recurrent neural network.