МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»

Методи штучного інтелекту в кібербезпеці

Практикум

*Рекомендовано Методичною радою КПІ ім. Ігоря Сікорського
як навчальний посібник для здобувачів ступеня доктора філософії за освітньою програмою «Кібербезпека»
спеціальності 125 «Кібербезпека»*

Київ
КПІ ім. Ігоря Сікорського
2022

Назва: Методи штучного інтелекту в кібербезпеці: Практикум [Електронний ресурс] : навч. посіб. для здобувачів спец. 125 «Кібербезпека» / КПІ ім. Ігоря Сікорського ; уклад.: І.В. Стьопочкіна – Електронні текстові дані (1 файл: 388,3 Кбайт). – Київ : КПІ ім. Ігоря Сікорського, 2022. – 18 с.

Гриф надано Методичною радою КПІ ім. Ігоря Сікорського (протокол № 4 від 07.04.2022 р.)
за поданням Вченої ради Фізико-технічного інституту (протокол № 21 від 20.12.2021 р.)

Електронне мережне навчальне видання

Методи штучного інтелекту в кібербезпеці

ПРАКТИКУМ

Укладач: *Стьопочкіна Ірина Валеріївна*, канд. техн. наук

Відповідальний редактор *Смирнов Сергій Анатолійович*, к.ф.-м.н., с.н.с.

Рецензент *Родіонов Андрій Миколайович, к.т.н., доцент кафедри інформаційної безпеки*

Методичні вказівки надають рекомендації щодо практичних занять з дисципліни “Методи штучного інтелекту в кібербезпеці”. У складі методичних вказівок наводиться план проведення заняття, надано основні етапи опанування матеріалу та завдання за відповідними темами. Надано рекомендації щодо виконання завдань, посилання на відповідні літературні джерела. Теми практичних занять пов’язані із темами, які розглядаються на лекційних заняттях.

ЗМІСТ

ВСТУП

Уміння застосовувати на практиці методи штучного інтелекту є дуже нагальним для задач кібербезпеки. Останнім часом більшість наукоємних задач кібербезпеки спираються саме на методи штучного інтелекту. Методичні вказівки дають можливість опанувати теоретичний матеріал дисципліни, сприяють підготовці до семестрового контролю. Надано переліки відповідних інформаційних джерел, за якими може здіснюватись підготовка здобувача.

В складі практичних занять присутні теми:

1. **Застосування методів штучного інтелекту для задач кібербезпеки.** Основна мета заняття – виконати огляд існуючих літературних джерел за темою методів штучного інтелекту в кібербезпеці, обрати та обгрунтувати вибір метода для теми власного дослідження. Усвідомити принципи виділення характеристик досліджуваного об'єкта, представлення їх у формі, придатній для сприйняття моделлю
2. **Принципи застосування глибинного навчання.** Основна мета – розглянути широко відомі види нейронних мереж, та сферу їх використання. Ознайомитись із принципами роботи нейронних мереж та їх конструювання. Усвідомити принципи оцінки точності результатів моделювання.
3. **Застосування методів штучного інтелекту для деяких класів задач кібербезпеки.** Основна мета – опанувати практичні прийоми, які використовуються при застосуванні підходів, що використовують моделі машинного навчання. Ознайомитись із методами покращення якості моделі.
4. **Застосування методів штучного інтелекту для деяких класів задач безпеки мереж.** Основна мета – ознайомитись із можливостями SAT/SMT розв’язників у застосуванні до безпеки мереж, та ознайомитись із методами класифікації трафіку для задач кібербезпеки.

1 ПРАКТИЧНЕ ЗАНЯТТЯ № 1. ЗАСТОСУВАННЯ МЕТОДІВ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ ЗАДАЧ КІБЕРБЕЗПЕКИ

1.1 Теоретичні відомості

Мета: виконати огляд та аналіз існуючих методів штучного інтелекту в кібербезпеці, обрати та обгрунтувати вибір метода, програмного забезпечення для теми власного дослідження. Усвідомити принципи виділення характеристик досліджуваного об'єкта, представлення їх у формі, придатній для сприйняття моделлю.

Теоретичні відомості до заняття наведені у навчальному посібнику [1, гл.1], графічному матеріалі у вигляді слайдів, які надаються здобувачу, а також джерелах [2-4].

1.2 План заняття

За наданим лекційним матеріалом та джерелами, вказаними впродовж лекції, надати відповідь на запитання:

1. Відмінність між класифікацією та кластеризацією.
2. Які алгоритми використовують для класифікації?
3. Які методи використовують для кластеризації?
4. В чому полягають особливості генеративних моделей та сфера їх застосування.
5. Для якої мети використовують методи зменшення розмірності, назвіть приклади.
6. Які типові задачі кібербезпеки можна розв’язувати із використанням методів машинного навчання?
7. Якими є переваги та недоліки регресійних моделей.
8. В чому полягає відмінність логістичної регресії від лінійної. Які задачі можна розв'язувати із використанням регресії.
9. Що таке дискримінантні функції? Яким чином проводиться їх вибір.

2. Сформулювати опис підзадачі власного дисертаційного дослідження, в якій можна застосувати методи машинного навчання.

Наприклад:

Варіант 1. Розпізнавання емоційної забарвленості тексту на основі заданої множини текстів соціальної мережі.

Варіант 2. Задача класифікації ШПЗ / задача генерації датасету, який характеризує ШПЗ (із використанням GAN).

Задачі можуть варіюватись в залежності від теми дисертаційного дослідження.

2.1. Описати перелік характеристик (features), які можуть бути вихідними даними для вирішення задачі.

2.2. Обгрунтувати вибір методу/методів штучного інтелекту, який може бути використаний для розв'язання поставленої задачі. Обгрунтування вибору відповідного методу підкріпити посиланнями на відповідні останні публікації в області дослідження (5-7 робіт).

Примітка: Для пошуку адекватних публікацій рекомендовано зареєструватись на www.scopus.com, а також у researchgate.net.

3. Здійснити пошук по моделям із відкритим кодом наприклад, на github. Здійснити огляд розроблених класифікаторів, які можуть бути використані для виконання експерименту (напр. WEKA classifier). Обрати рішення, яке може бути придатним для здійснення обчислювального експерименту.

2 ПРАКТИЧНЕ ЗАНЯТТЯ № 2. ПРИНЦИПИ ЗАСТОСУВАННЯ ГЛИБИННОГО НАВЧАННЯ.

2.1 Теоретичні відомості

Мета: розглянути широко відомі види нейронних мереж, та сферу їх використання. Ознайомитись із принципами роботи нейронних мереж та їх конструювання, виконати порівняльний аналіз та існуючі інструментальні засоби. Усвідомити принципи оцінки точності результатів моделювання.

Теоретичні відомості до заняття наведені у навчальному посібнику [1, гл.2], графічному матеріалі у вигляді слайдів, які надаються здобувачу, а також джерелах [5,6].

2.2 План заняття

1. За наданим лекційним матеріалом та джерелами, вказаними впродовж лекції, надати відповідь на запитання:

1. Який принцип асоціативного запам'ятовування у розподілених системах типу Хопфілда-Хебба?
2. Перелічіть недоліки двошарового персептрона та системи Хопфілда-Хебба.
3. Які співвідношення для корекції вагових коефіцієнтів використовують для двошарового персептрона?
4. Які співвідношення для корекції вагових коефіцієнтів використовують для багатошарового (узагальненого) персептрона?
5. В яких задачах використовують моделі Больцмана? В чому їх відмінність від персептронів та систем Хопфілда-Хебба?
6. Яким чином використовується ROC-крива для одержання уявлення про успішність застосування методу? Як вона будується?
7. Якими є недоліки критеріїв оцінки класифікатора – точності та повноти. Чому F- міра є більш повним критерієм.
8. Коли варто використовувати confusion matrix для оцінки точності та повноти класифікатора? Як вона будується?

9. В чому полягає явище перенавчання та як його визначити за Learning Curve?
10. Чим ми керуємось, коли визначаємо розмір навчальної виборки?

2. Проаналізувати різні типи моделей із глибинним навчанням. Які підходи до побудови шарів та взаємодій між ними використовуються?

Розглянути:

Рекурентні мережі (RNN), зокрема:

- LSTM https://uk.wikipedia.org/wiki/Довга_короткочасна_пам%27ять ;
- BLSTM - розпізнавання емоцій по голосу;
- CNN (згорткові мережі) https://uk.wikipedia.org/wiki/Згорткова_нейронна_мережа#Розрізнювання_ознак ;
- FNN (мережі прямого розповсюдження) https://en.wikipedia.org/wiki/Feedforward_neural_network

Скласти таблицю порівняльних властивостей цих мереж (табл.1). Окреслити сферу застосування.

Таблиця 1. Порівняльні властивості мереж

Тип мережі	Кількість шарів	Взаємодії між шарами (передаткова функція, можливість циклічних взаємодій тощо)	Сфера застосування	Недоліки	Переваги

3. Ознайомитись із можливостями https://www.tensorflow.org/ по реалізації рекурентних та згорткових мереж.

П. 1,2 оформити у виді звіту, п.3 – зробити висновки по застосовності можливостей https://www.tensorflow.org/ до завдань дисертаційного дослідження.

3 ПРАКТИЧНЕ ЗАНЯТТЯ № 3. ЗАСТОСУВАННЯ МЕТОДІВ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ ДЕЯКИХ КЛАСІВ ЗАДАЧ КІБЕРБЕЗПЕКИ

3.1 Теоретичні відомості

Мета: опанувати практичні прийоми, які використовуються при застосуванні підходів, що використовують моделі машинного навчання. Ознайомитись із методами покращення якості моделі, розробити відповідний алгоритм ансамблювання класифікаторів.

Теоретичні відомості до заняття наведені у навчальному посібнику [1, гл.3, 4], графічному матеріалі у вигляді слайдів, які надаються здобувачу, а також джерелах [5,6, 10, 17].

3.2 План заняття

1. За наданим лекційним матеріалом та джерелами, вказаними впродовж лекції, надати відповідь на запитання:

11. Які основні етапи треба виконати, щоби представити вихідні текстові дані, у виді, придатному для обробки математичною моделлю у виді нейронної мережі. (в якості вихідних даних може бути а) код б) текстові повідомлення).
12. Для чого використовують структури типу абстрактного синтаксичного дерева при аналізі вихідних даних.
13. Яким чином необхідно підготувати датасет для навчання, щоби якість навчання була якомога вищою?
14. Які вимоги задачі треба врахувати при підборі типу моделі (зокрема, типу нейронної мережі)?
15. В чому основна відмінність LSTM мереж від BLSTM. Наведіть приклади задач для яких застосовуються перші та другі мережі.
16. Які групи характеристик можна виділити для виявлення фішингових сайтів?

17. Які методи можна використати для того, щоби виявити найбільш значущі характеристики і відсіяти другорядні. Чи завжди відсіювання другорядних характеристик відбивається на точності класифікації?
18. Яким чином можна комбінувати класифікатори, які дають не дуже хороші результати, забезпечуючи покращення точності класифікації.
19. Що таке boosting, яким є принцип цього підходу https://towardsdatascience.com/understanding-adaboost-2f94f22d5bfe ?

2. Проаналізуйте алгоритми бустингу (LPBoost, XGBoost https://github.com/dmlc/xgboost/releases/tag/v1.2.1 чи інші http://www.machinelearning.ru/wiki/images/9/9a/fonarev.overview_of_boosting_methods.pdf). Чи можуть вони бути застосовані до вашої задачі дисертаційного дослідження?

3. Розглянути можливості методу регуляризації моделей глибинного навчання за допомогою методу dropout.

4. Запропонуйте алгоритм, яка комбінує декілька класифікаторів, та визначте яким чином можна регулювати якість роботи такого ансамблю в контексті вашої задачі (псевдокод варіанту алгоритму є в http://www.machinelearning.ru/wiki/images/9/9a/fonarev.overview_of_boosting_methods.pdf). Чи можуть вони бути застосовані до вашої задачі дисертаційного).

Наприклад, загальний алгоритм двійкової класифікації може виглядати таким чином:

1. Формуємо великий набір прикладів.
2. Ініціюємо ваги для прикладів тренувального набору, нормалізуємо їх.
3. Робимо T ітерацій:

3.1. Для прикладів із тренувальної частини набору тренуємо класифікатор, обчислюємо похибку тренувального етапу на цих прикладах.

3.2. Обираємо класифікатор із найменшою похибкою розпізнавання.

3.3. Обновляємо ваги тренувальних прикладів: збільшуємо, якщо приклад класифіковано невірно, і зменшуємо, якщо вірно.

4. Формуємо посилений класифікатор як лінійну комбінацію N класифікаторів – коефіцієнт класифікатора більше, якщо похибка тренування менше.

4 ПРАКТИЧНЕ ЗАНЯТТЯ № 4. ЗАСТОСУВАННЯ МЕТОДІВ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ ДЕЯКИХ ЗАДАЧ БЕЗПЕКИ МЕРЕЖ

4.1 Теоретичні відомості

Мета: ознайомитись із можливостями SAT/SMT розв’язників у застосуванні до безпеки мереж, та методами класифікації трафіку для задач кібербезпеки.

Теоретичні відомості до заняття наведені у навчальному посібнику [1, гл.4-7], графічному матеріалі у вигляді слайдів, які надаються здобувачу, а також джерелах [15, 18-21].

4.2 План заняття

1. За наданим лекційним матеріалом та онлайн-джерелами, вказаними впродовж лекції, надати відповідь на запитання:

1. Які задачі аналізу безпеки комп’ютерних систем та мереж можна розв’язати методами штучного інтелекту. Наведіть приклади.

2. Які методи доречно застосовувати для а) двокласової б) мультикласової класифікації трафіку пристроїв IoT. Яка мета такої класифікації в контексті задач кібербезпеки?
3. Які характеристики можна використовувати для класифікації трафіку пристроїв IoT.
4. Якими є переваги алгоритму Random Forest, в яких задачах кібербезпеки доречне його використання.
5. В яких задачах кібербезпеки мереж можливе застосування SMT – розв`язників.

2. Проаналізуйте можливості розв'язника Z3 як інструменту підтримки експертних рішень https://theory.stanford.edu/~nikolaj/programmingz3.html . Чи можуть вони бути застосовані до вашої задачі дисертаційного дослідження? Яким чином можна здійснити перехід від комп'ютерних мереж до графів, а потім до булевих виразів, які сприймаються розв'язником?

3. Ознайомтесь із можливостями бібліотеки scikit-learn в частині класифікатора Random Forest. На що впливають його основні параметри, як їх рекомендовано задавати.

ПЕРЕЛІК ТЕМ, РЕКОМЕНДОВАНИХ ДО ІНДИВІДУАЛЬНИХ ЗАНЯТЬ

Окрім тем, розглянутих в [1], рекомендуються наступні теми із використанням методів штучного інтелекту:

1. Генерація цільового голосу для тестування на проникнення методами соціальної інженерії [26].
2. Виявлення фейкових аккаунтів в соціальних мережах [1,28].
3. Виявлення фішингових сайтів [28].
4. Виявлення тихих зворотних серверів при зловмисному перехопленні параметрів автентифікації [25].

5. Надання рекомендацій по обробці ризиків [7].
6. Визначення дієвості політики безпеки на основі моделей регресії [6,7].
7. Генерація текстів природної мови для автоматизованого компонування листів при тестуванні на проникнення методами соціальної інженерії [24,27].
8. Виявлення слабкостей та вразливостей програмного коду [13,14, 32].
9. Алгоритми самотестування та самовідновлення інтелектуальних систем [22,23].
10. Виявлення небезпечних ознак в відеозображеннях в задачах забезпечення організаційної безпеки [5].
11. Методи виявлення ШПЗ за ознаками опкодів, PE-заголовків, системних викликів, особливостей бінарних файлів [5,6,29-31].

ЛІТЕРАТУРА

1. Методи штучного інтелекту в кібербезпеці [Електронний ресурс] : навч. посіб. для здобувачів спец. 125 «Кібербезпека» / КПІ ім. Ігоря Сікорського ; уклад.: І.В. Стьопочкіна, О.М. Новіков. – Електронні текстові дані (1 файл: X,XX Мбайт). – Київ : КПІ ім. Ігоря Сікорського, 2022 – 81 с.
2. S. Varshney et al, Review of various Artificial Intelligence Techniques and its applications, 2019. IOP Conf. Ser.: Mater. Sci. Eng. 594 012023, DOI: 10.1088/1757-899X/594/1/012023.
3. G. Belani, The Use of Artificial Intelligence in Cybersecurity: A Review. URL: https://www.computer.org/publications/tech-news/trends/the-use-of-artificial-intelligence-in-cybersecurity.
4. S. Ray, 7 Regression Techniques you should know! URL: https://www.analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/
5. C.M.Bishop, Pattern recognition and Machine learning, 2006. Springer.
6. K.P.Murphy, Machine Learning: A Probabilistic Perspective, 2012. MIT Press.
7. В. Кисіль, Політика безпеки для промислового Інтернету речей та оцінка її дієвості, 2019. Магістерська дисертація. URL: https://ela.kpi.ua/bitstream/123456789/31393/1/Kysil_magistr.pdf
8. В. Кожокар, І. Стьопочкіна, Структурні закономірності еволюціонування метаморфного шкідливого ПЗ// Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні.-2017, №1(33).-С.52-57.
9. Кожокар В.Ю., Стьопочкіна І.В. Виявлення спільних закономірностей у зразках метаморфного шкідливого ПЗ на основі статичних характеристик. XV Всеукраїнська науково-практична конференція студентів, аспірантів та молодих вчених “Теоретичні і прикладні проблеми фізики, математики та інформатики” 25-27 травня 2017, т.2.
10. J.Brownlee, A Gentle Introduction to Dropout for Regularizing Deep Neural Networks, 2018. URL: https://machinelearningmastery.com/dropout-for-regularizing-deep-neural-networks/
11. Learning curve. URL: https://www.ritchieng.com/machinelearning-learning-curve/
12. A. Gonfalonieri, How to Build A Data Set For Your Machine Learning Project, 2019. URL: https://towardsdatascience.com/how-to-build-a-data-set-for-your-machine-learning-project-5b3b871881ac

13. А. Черноусов, Метод автоматичного виявлення помилок безпеки в програмному забезпеченні на основі глибинного навчання, 2019. Магістерська дисертація. URL:https://ela.kpi.ua/handle/123456789/31602
14. Clang 15.0.0git documentation. URL: https://clang.llvm.org/docs/LibASTMatchersTutorial.html
15. В. Мельник, Класифікація пристроїв Інтернету речей з використанням методів машинного навчання, 2019. Магістерська дисертація. URL: https://ela.kpi.ua/handle/123456789/31387
16. Sklearn.ensemble.RandomForestClassifier. URL: https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html
17. Tuning the hyper-parameters of an estimator. URL: https://scikit-learn.org/stable/modules/grid_search.html
18. SAT/SMT by example. URL: https://sat-smt.codes/SAT_SMT_by_example.pdf
19. S. Mitra, Tutorial 1: Modern SMT Solvers and Verification. URL: https://www.iitg.ac.in/pbhaduri/GIAN-CPS/Doc/Tut-1.pdf
20. К. Рішко, Застосування SAT/SMT- розв'язників в задачах кібербезпеки, 2020. Магістерська дисертація.
21. SMT-LIB. The Satisfability Modulo Theories Library. https://smtlib.cs.uiowa.edu
22. О. Фокін, Технології самовідновлення для розподілених систем на базі штучного інтелекту, 2019. Магістерська дисертація. URL: https://ela.kpi.ua/bitstream/123456789/31604/1/Fokin_magistr.pdf
23. J.Zhang. U-Net. Line-by-line explaination, 2019. URL: https://towardsdatascience.com/unet-line-by-line-explanation-9b191c76baf5
24. EleutherAI. URL: https://www.eleuther.ai
25. Vlasenko A., Stopochkina I., Ilin M. Methods of counteraction of bypassing two-factor authentication using reverse proxy //Theoretical and Applied Cybersecurity Vol. 3 No. 1 (2021): Theoretical And Applied Cybersecurity.
26. Маковецький А.О., Стьопочкіна І.В. Застосування технологій deepfake в задачі тестування на проникнення // Матеріали Всеукраїнської науково-практичної конференції "Теоретичні і прикладні проблеми фізики, математики та інформатики", 13-14 травня 2021 р., м Київ, с. 361-364.
27. С. В. Іванченко, І. В. Стьопочкіна. Методи автоматизованої генерації фішингових повідомлень для тестування на проникнення //Матеріали Всеукраїнської науково-практичної конференції "Теоретичні і прикладні проблеми фізики, математики та інформатики", 12-13 травня 2020, - с.257-260.

28. С. Тернопольська. Розпізнавання фішингових сайтів з використанням методів машинного навчання, 2019. Бакалаврська робота. URL: https://ela.kpi.ua/bitstream/123456789/31430/1/Ternopolska_bakalavr.pdf
29. Detection Metamorphic Virus Detection in Portable Executables Using Opcodes Statistical Feature, 2014. URL: https://arxiv.org/ftp/arxiv/papers/1104/1104.3229.pdf
30. Babak B. R. , Metamorphic Virus Detection in Portable Executables Using Opcodes Statistical Feature, 2015. URL: https://arxiv.org/abs/1104.3229.
31. Konstantinou E. Metamorphic Virus: Analysis and Detection, 2015. URL: https://www.ma.rhul.ac.uk/static/techrep/2008/RHUL-MA-2008-02.pdf
32. Anders Møller and Michael I. Schwartzbach. Static Program Analysis. URL: https://cs.au.dk/~amoeller/spa/