Силабус освітнього компонента

Програма навчальної дисципліни

Machine Learning

Шифр та назва спеціальності
F5 – Кібербезпека та захист інформації

Інститут
ННІ комп`ютерних наук та інформаційних технологій

Спеціалізація

Кафедра
Кібербезпеки (328)

Освітня програма
Кібербезпека

Тип дисципліни
Професійна підготовка, Вибіркова

Рівень освіти
Другий (магістерський)

Форма навчання
Денна, заочна

Семестр
2

Мова викладання
Українська, англійська

Викладачі, розробники

ЄВСЕЄВ Сергій Петрович

serhii.yevseiev@khpi.edu.ua
Доктор технічних наук, професор, завідувач кафедри кібербезпеки НТУ «ХПІ».

Лауреат національної премії імені Бориса Патона.
Кількість наукових публікацій: понад 351, з них патентів на корисну модель 44, 11 монографій, з яких 10 колективних монографій, 31 навчальний посібник, з яких 4 з грифом Міністерства освіти і науки України, 149 статтей у закордонних виданнях та фахових виданнях України, з них 61 у наукометричній базі Scopus. Гарант освітньо-професійної програми "Кібербезпека" першого (бакалаврського) рівня вищої освіти. Провідний лектор з дисциплін: Математичні основи криптології, Комп`ютерні мережі, Основи криптографічного захисту, Методи криптоаналізу, Основи кібербезпеки, Сучасні проблеми постквантової криптографії, Основи наукових досліджень, Цифрова криміналістика, Етичний хакінг у студентів бакалавріата та магістратури, Методологія наукової та педагогічної діяльності в науках кіберзахисту для аспірантів.

Детальніше про викладача на сайті кафедри

Загальна інформація

Анотація

Навчальна програма дисципліни "Maching Learning" визначає зміст і обсяг знань, необхідних для фахівця з кібербезпеки. Дисципліна обіймає проблематику вивчення сучасного стану технологій машинного навчання, що використовуються для формалізації та обробки даних в технологіях функціонування систем безпеки, вивчення сучасних програмних засобів машинної обробки даних, технологій їх проектування, реалізації, налагодження і дослідження. Машинне навчання застосовують в ряді обчислювальних задач, в яких розробка та програмування явних алгоритмів з доброю продуктивністю є складною або нездійсненною задачею. Для практичного засвоєння навчальних матеріалів ряд тем дисципліни поглиблено вивчається на лабораторних заняттях.

Мета та цілі дисципліни

Метою вивчення дисципліни є формування у майбутніх фахівців знань та вмінь застосування сучасних методів та засобів розробки, дослідження та використання сучасних технологій обробки даних для вирішення задач класифікації, регресійного аналізу, прогнозування та прийняття рішень.

Формат занять

Лекції, лабораторні заняття, самостійна робота, консультації. Підсумковий контроль – диференційований залік.

Компетентності

КФ1. Здатність обґрунтовано застосовувати, інтегрувати, розробляти та удосконалювати сучасні інформаційні технології, фізичні та математичні моделі, а також технології створення та використання прикладного і спеціалізованого програмного забезпечення для вирішення професійних задач у сфері інформаційної безпеки та/або кібербезпеки.
КФ3. Здатність досліджувати, розробляти і супроводжувати методи та засоби інформаційної безпеки та/або кібербезпеки на об'єктах інформаційної діяльності та критичної інфраструктури.
КФ4. Здатність аналізувати, розробляти і супроводжувати систему управління інформаційною безпекою та/або кібербезпекою організації, формувати стратегію і політики інформаційної безпеки з урахуванням вітчизняних і міжнародних стандартів та вимог.
КФ5. Здатність до дослідження, системного аналізу та забезпечення безперервності бізнес/операційних процесів з метою визначення вразливостей інформаційних систем та ресурсів, аналізу ризиків та визначення оцінки їх впливу у відповідності до встановленої стратегії і політики інформаційної безпеки та/або кібербезпеки організації.
КФ6. Здатність аналізувати, контролювати та забезпечувати систему управління доступом до інформаційних ресурсів згідно встановленої стратегії і політики інформаційної безпеки та/або кібербезпеки організації.
КФ7. Здатність досліджувати, розробляти та впроваджувати методи і заходи протидії кіберінцидентам, здійснювати процедури управління, контролю та розслідування, а також надавати рекомендації щодо попередження та аналізу кіберінцидентів в цілому.

Результати навчання

РН2. Інтегрувати фундаментальні та спеціальні знання для розв'язування складних задач інформаційної безпеки та/або кібербезпеки у широких або мультидисциплінарних контекстах.
РН3. Провадити дослідницьку та/або інноваційну діяльність в сфері інформаційної безпеки та/або кібербезпеки, а також в сфері технічного та криптографічного захисту інформації у кіберпросторі.
РН4. Застосовувати, інтегрувати, розробляти, впроваджувати та удосконалювати сучасні інформаційні технології, фізичні та математичні методи і моделі в сфері інформаційної безпеки та/або кібербезпеки.
РН5. Критично осмислювати проблеми інформаційної безпеки та/або кібербезпеки, у тому числі на міжгалузевому та міждисциплінарному рівні, зокрема на основі розуміння нових результатів

інженерних і фізико-математичних наук, а також розвитку технологій створення та використання спеціалізованого програмного забезпечення.
РН6. Аналізувати та оцінювати захищеність систем, комплексів та засобів кіберзахисту, технології створення та використання спеціалізованого програмного забезпечення.
РН7. Обґрунтовувати використання, впроваджувати та аналізувати кращі світові стандарти, практики з метою розв'язання складних задач професійної діяльності в галузі інформаційної безпеки та/або кібербезпеки.
РН8. Досліджувати, розробляти і супроводжувати системи та засоби інформаційної безпеки та/або кібербезпеки на об'єктах інформаційної діяльності та критичної інфраструктури.
РН9. Аналізувати, розробляти і супроводжувати систему управління інформаційною безпекою та/або кібербезпекою організації на базі стратегії і політики інформаційної безпеки.
РН10. Забезпечувати безперервність бізнес/операційних процесів, а також виявляти уразливості інформаційних систем та ресурсів, аналізувати та оцінювати ризики для інформаційної безпеки та/або кібербезпеки організації.
РН11. Аналізувати, контролювати та забезпечувати ефективне функціонування системи управління доступом до інформаційних ресурсів відповідно до встановлених стратегії і політики інформаційної безпеки та/або кібербезпеки організації.
РН12. Досліджувати, розробляти та впроваджувати методи і заходи протидії кіберінцидентам, здійснювати процедури управління, контролю та розслідування, а також надавати рекомендації щодо попередження та аналізу кіберінцидентів в цілому.
РН15. Зрозуміло і недвозначно доносити власні висновки з проблем інформаційної безпеки та/або кібербезпеки, а також знання та пояснення, що їх обґрунтовують до персоналу, партнерів та інших осіб.
РН16. Приймати обґрунтовані рішення з організаційно-технічних питань інформаційної безпеки та/або кібербезпеки у складних і непередбачуваних умовах, у тому числі із застосуванням сучасних методів та засобів оптимізації, прогнозування та прийняття рішень.
РН17. Мати навички автономного і самостійного навчання у сфері інформаційної безпеки та/або кібербезпеки і дотичних галузей знань, аналізувати власні освітні потреби та об'єктивно оцінювати результати навчання.
РН19. Обирати, аналізувати і розробляти придатні типові аналітичні, розрахункові та експериментальні методи кіберзахисту, розробляти, реалізовувати та супроводжувати проекти з захисту інформації у кіберпросторі, інноваційної діяльності та захисту інтелектуальної власності.
РН20. Ставити та вирішувати складні інженерно-прикладні та наукові задачі інформаційної безпеки та/або кібербезпеки з урахуванням вимог вітчизняних та світових стандартів та кращих практик.
РН21. Використовувати методи натурного, фізичного і комп'ютерного моделювання для дослідження процесів, які стосуються інформаційної безпеки та/або кібербезпеки.
РН22. Планувати та виконувати експериментальні і теоретичні дослідження, висувати і перевіряти гіпотези, обирати для цього придатні методи та інструменти, здійснювати статистичну обробку даних, оцінювати достовірність результатів досліджень, аргументувати висновки.
РН23. Обґрунтовувати вибір програмного забезпечення, устаткування та інструментів, інженерних технологій і процесів, а також обмежень щодо них в галузі інформаційної безпеки та/або кібербезпеки на основі сучасних знань у суміжних галузях, наукової, технічної та довідкової літератури та іншої доступної інформації.

Обсяг дисципліни

Загальний обсяг дисципліни 120 год. (4 кредити ECTS): лекції – 16 год., лабораторні роботи – 32 год., самостійна робота – 72 год.

Передумови вивчення дисципліни (пререквізити)

Вища математика.

Особливості дисципліни, методи та технології навчання

В ході викладання дисципліни викладачем застосовуються пояснювально-ілюстративний (інформаційно-рецептивний) та репродуктивний методи навчання. В якості методів викладання, які направлені на активізацію та стимулювання навчально-пізнавальної діяльності здобувачів, застосовуються презентації, бесіди, індивідуальні групові проєкти, майстер-класи.

Програма навчальної дисципліни

Навчальні заняття

Лекції

Теми лекцій	Кількість годин
Тема 1. Вступ до машиного навчання. Позначення та визначення. Що таке машинне навчання. Типи навчання. Як працює навчання із учителем. Чому модель здатна працювати з новими даними? Позначення. Випадкова величина. Незміщені оцінки. Правило Байєса. Оцінка параметрів. Параметри та гіперпараметри. Класифікація та регресія. Навчання на основі моделей та на основі прикладів. Поверхневе та глибоке навчання.	2
Тема 2. Фундаментальні алгоритми. Анатомія алгоритмів навчання. Лінійна регресія. Логістична регресія. Навчання дерева рішень. Спосіб опорних векторів. Метод k найближчих сусідів. Будівельні блоки алгоритмів навчання. Градієнтний спуск. Як працюють інженери, котрі займаються машинним навчанням. Особливості алгоритмів навчання.	2
Тема 3. Практичні засади. Проектування ознак. Вибір алгоритму навчання. Три набори. Недонавчання та перенавчання. Регуляризація. Оцінка ефективності моделі. Налаштування гіперпараметрів.	2
Тема 4. Нейронні мережі та глибоке навчання. Нейронні мережі. Глибоке навчання.	2
Тема 5. Проблеми та рішення. Ядерна регресія. Багатокласова класифікація. Однокласова класифікація. Класифікація з багатьма мітками. Навчання ансамблю. Навчання маркування послідовностей. Навчання перетворення послідовностей у послідовності. Активне навчання. Навчання із частковим залученням вчителя. Навчання з першого разу. Навчання без підготовки.	2
Тема 6. Просунуті методики. Робота з незбалансованими наборами даних. Об'єднання моделей. Навчання нейронних мереж. Просунута регуляризація. Обробка кількох входів. Обробка кількох виходів. Перенесення навчання. Ефективність алгоритмів.	2
Тема 7. Навчання без учителя. Оцінка густини. Кластеризація. Скорочення розмірності. Виявлення аномалій.	2
Тема 8. Інші форми навчання. Визначення метрик. Визначення рангу. Навчання робити рекомендації. Самонавчання з учителем: вкладення слів.	2
Загальна кількість годин	**16**

Лабораторні заняття

Контрольні роботи

Теми лабораторних занять	Кількість годин	Вагові коефіцієнти a
Тема 1. Основи машинного навчання, базові алгоритми та практичні принципи побудови моделей. Вступ до машинного навчання: типи навчання, ключові позначення та поняття; лінійна й логістична регресія, дерева рішень, SVM, k-NN; градієнтний спуск і будова алгоритмів; проектування ознак, недонавчання/перенавчання, регуляризація, оцінка ефективності та налаштування гіперпараметрів.	4	0,1
Тема 2. Важливість та нормалізація ознак. Вивчення методів масштабування та нормалізації (Min-Max, StandardScaler). Оцінка важливості ознак, її вплив на якість моделі та процес відбору релевантних параметрів.	4	0,1
Тема 2. Глибоке навчання, спеціалізовані методики та розширені підходи до моделювання. Нейронні мережі та глибоке навчання; багатокласові, однокласові та багатомітокові задачі, ансамблеві методи, навчання послідовностей; активне та частково контрольоване навчання, робота з незбалансованими даними; перенесення навчання, просунута регуляризація, навчання без учителя — кластеризація, скорочення розмірності, аномалії; рекомендаційні системи та визначення метрик.	4	0,1
Тема 3. Вибір числа сусідів, вибір метрики. Практичний підбір оптимального числа k для методу k-NN. Вибір і порівняння метрик відстані (Euclidean, Manhattan, Minkowski) та їх вплив на результат.	4	0,1
Тема 4. Опорні об'єкти. Дослідження роботи алгоритму SVM, вибір ядра, визначення опорних об'єктів. Аналіз того, як гіперплощина розділяє класи та як змінюються опорні вектори при налаштуванні параметрів.	4	0,1
Тема 5. Аналіз текстів. Передобробка текстових даних: токенізація, очищення, лематизація. Побудова моделей на основі Bag-of-Words, TF-IDF. Класифікація текстів за допомогою базових алгоритмів машинного навчання.	4	0,1
Тема 6. Логістична регресія. Лінійна регресія. Побудова моделей регресії, підбір параметрів, оцінка якості (MAE, MSE, R^2, accuracy). Практичне порівняння регресійних моделей на різних наборах даних.	4	0,1
Тема 7. Розмір випадкового лісу. Робота з алгоритмом Random Forest: підбір кількості дерев, глибини дерев, аналіз важливості ознак. Оцінка впливу гіперпараметрів на точність і стійкість моделі.	4	0,1
Тема 8. Метод найшвидшого бустингу над вирішальними деревами. Вивчення градієнтного бустингу (XGBoost, LightGBM чи CatBoost). Налаштування параметрів, оптимізація швидкодії, оцінка продуктивності моделі на практичному завданні.	4	0,1
Загальна кількість годин	**32**	$\sum_{i=1}^{n} a_i = 0{,}8$

Теми контрольних робіт	Вагові коефіцієнти b
Загалом	$\sum_{i=1}^{m} b_i = 0{,}2$

Самостійна робота

До самостійної роботи відноситься самостійне опрацювання теоретичного матеріалу

Опрацювання теоретичного матеріалу

Теми для самостійного вивчення	Кількість годин
Тема 1 Регуляризація у глибокому навчанні. L1/L2, Dropout, Batch Normalization, Early Stopping — коли і як застосовуються.	8

Тема 2. Оптимізатори в нейронних мережах. SGD, Momentum, Adam, RMSProp — порівняння та практичні рекомендації.	8
Тема 3. Інтерпретованість моделей (Explainable AI). SHAP, LIME, Partial Dependence Plots — пояснення рішень моделей.	7
Тема 4. Робота з незбалансованими даними. Oversampling, undersampling, SMOTE, AUC/ROC та інші метрики.	7
Тема 5. Рекомендаційні системи. Колаборативна фільтрація, матрична факторизація, контентні рекомендації.	7
Тема 6. Глибокі нейронні мережі для обробки зображень. CNN, ResNet, Transfer Learning, Data Augmentation.	7
Тема 7. Моделі для обробки природної мови (NLP). Word2Vec, GloVe, Transformer, BERT — сучасні методи роботи з текстами.	7
Тема 8. Байєсівські моделі та ймовірнісне навчання. Наївний Байєс, апріорні/апостеріорні розподіли, MCMC.	7
Тема 9. Розгортання моделей ML у продакшн. Docker, REST API, CI/CD, моніторинг моделей, MLOps.	7
Тема 10. Підсилення навчання (Reinforcement Learning). Агенти, функції винагороди, Q-learning, політики, середовище навчання.	7
Загальна кількість годин	**72**

Неформальна освіта

Здобувач має можливість перезарахувати окремі теми або курс шляхом: проходження професійних курсів чи тренінгів, онлайн-освіти, професійних стажувань, у сфері, що відповідає навчальним цілям дисципліни.
Для зарахування необхідно надати: сертифікат (електронний або друкований) про проходження курсу/стажування, опис програми тренінгу із зазначенням змісту тем, обсягу та тривалості.

Рекомендовані курси, тренінги, стажування

1. Онлайн-курс CISCO «AI Fundamentals with IBM SkillsBuild, Learning Collection: AI with Intel» https://www.netacad.com/catalogs/learn?category=course.

Література, навчальні матеріали та інформаційні ресурси

Основна література

1. Bishop, Christopher M. Pattern Recognition and Machine Learning. Springer;
1st ed. 2006. Corr.; 2nd printing ed. 2007. [Електронний ресурс]. – Режим доступу:
https://www.microsoft.com/en-us/research/uploads/prod/2006/01/Bishop-Pattern-Recognition-and-Machine-Learning-2006.pdf
2. Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. The Elements of Statistical Learning. Springer, 2009. [Електронний ресурс]. – Режим доступу:
https://www.sas.upenn.edu/~fdiebold/NoHesitations/BookAdvanced.pdf
3. Gelman, Andrew, and Jennifer Hill. Data Analysis Using Regression and Multilevel/
Hierarchical Models. Cambridge University Press, 2006. [Електронний ресурс]. – Режим доступу:
https://moodle2.units.it/pluginfile.php/706395/mod_resource/content/1/gelman%20hill%202007.pdf
4. Jones, Owen, Robert Maillardet, and Andrew Robinson. Introduction to Scientific Programming and Simulation Using R. Chapman and Hall, 2009. [Електронний ресурс]. – Режим доступу:
https://nyu-cdsc.github.io/learningr/assets/simulation.pdf

5. Segaran, Toby. Programming Collective Intelligence: Building Smart Web 2.0 Applications. O'Reilly Media, 2007. [Електронний ресурс]. – Режим доступу: https://axon.cs.byu.edu/~martinez/classes/778/Papers/GP.pdf

Додаткова література

6. A. Rukhin, and J. Soto. A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications. NIST Special Publication 800-22, 09.2000, 164 p. [Електронний ресурс]. – Режим доступу: https://nvlpubs.nist.gov/nistpubs/legacy/sp/nistspecialpublication800-22r1a.pdf

Система оцінювання

Підсумкова оцінка з освітнього компонента визначається відповідальним лектором за темами, видами занять, тощо у відповідності до силабусу і є інтегральною оцінкою результатів усіх вид навчальної діяльності здобувача вищої освіти. Підсумкова оцінка повинна відображати всі оцінки за складовими навчального процесу з урахуванням їх вагових показників k:

Поточний контроль (практичні, семінарські, лабораторні заняття), k_1	Контрольні роботи (за наявності), k_2	Індивідуальне завдання (за наявності), k_3	Підсумковий контроль (для ОК з іспитом), k_4
0,8	0,2		

Сума коефіцієнтів повинна складати одиницю: $k_1 + k_2 + k_3 + k_4 = 1$. Підбір вагових коефіцієнтів підсумкової оцінки здійснює розробник курсу.
Розрахунок підсумкової оцінки проводиться за формулою:

$$О = П \cdot k_1 + K \cdot k_2 + І \cdot k_3 + Пк \cdot k_4$$

де: П – середньозважена середня оцінка за поточний контроль
І – оцінка за виконання індивідуального завдання
K – середньозважена оцінка за контрольні роботи
Пк – оцінка за підсумковий контроль

$$П = \frac{П_1 \cdot a_1 + П_2 \cdot a_2 + ... + П_n \cdot a_n}{\sum_{i=1}^{n} a_i}$$

де: a_i - ваговий коефіцієнт за кожне практичне (семінарське) або лабораторне заняття.

$$K = \frac{K_1 \cdot b_1 + K_2 \cdot b_2 + ... + K_m \cdot b_m}{\sum_{i=1}^{m} b_i}$$

де: b_i - ваговий коефіцієнт за кожну контрольну роботу.

Поточні оцінки за кожну складову (П, K, І,...) виставляються за 100-бальною шкалою згідно з положенням «Про критерії та систему оцінювання знань та вмінь і про рейтинг здобувачів вищої освіти» НТУ «ХПІ».

Підсумкова оцінка виставляється відповідно до розрахованої О з округленням до найближчого цілого числа в більшу сторону.

Шкала оцінювання

Сума балів	Національна оцінка	ECTS
90–100	Відмінно	A
82–89	Добре	B
75–81	Добре	C
64–74	Задовільно	D
60–63	Задовільно	E
35–59	Незадовільно (потрібне додаткове вивчення)	FX
1–34	Незадовільно	F

(потрібне повторне вивчення)

Норми академічної етики і політика курсу

Здобувач вищої освіти повинен дотримуватися «Кодексу етики академічних взаємовідносин та доброчесності НТУ «ХПІ»: виявляти дисциплінованість, вихованість, доброзичливість, чесність, відповідальність. Конфліктні ситуації повинні відкрито обговорюватися в навчальних групах з викладачем, а при неможливості вирішення конфлікту – доводитися до відома співробітників дирекції інституту.
Нормативно-правове забезпечення впровадження принципів академічної доброчесності НТУ «ХПІ» розміщено на сайті: http://blogs.kpi.kharkov.ua/v2/nv/akademichna-dobrochesnist/

Погодження

Силабус погоджено	30.08.2025		**Завідувач кафедри** Сергій ЄВСЕЄВ
	30.08.2025		**Гарант ОП** Ольга КОРОЛЬ